LAW, CULTURE AND THE HUMANITIES

MINI SYMPOSIUM: SCIENCE, SCIENCE FICTION AND INTERNATIONAL LAW

Law, Culture and the Humanities
2018, Vol. 14(1) 45–82

DOI: 10.1177/1743872113509443
journals.sagepub.com/home/lch
SAGE

Evitable Conflicts, Inevitable Technologies? The Science and Fiction of Robotic Warfare and IHL

Ian Kerr
University of Ottawa, Canada

Katie Szilagyi
McCarthy Tétrault LLP, Canada

Abstract

This article contributes to a special symposium on science fiction and international law, examining the blurry lines between science and fiction in the policy discussions concerning the military use of lethal autonomous robots. In response to projects that attempt to build military robots that comport with international humanitarian law [IHL], we investigate whether and how the introduction of lethal autonomous robots might skew international humanitarian norms. Although IHL purports to be a technologically-neutral approach to calculating a proportionate, discriminate, and militarily necessary response, we contend that it permits a deterministic mode of thinking, expanding the scope of that which is perceived of as "necessary" once the technology is adopted. Consequently, we argue, even if lethal autonomous robots comport with IHL, they will operate as a force multiplier of military necessity, thus skewing the proportionality metric and amplifying new forms of destructive, lethal force.

Keywords

International humanitarian law, science fiction, autonomous robots, proportionality, discrimination, military necessity, roboethics, technology

In 1967—the same year that IBM invented the floppy disk, the US bombed Hanoi, the USSR ratified a treaty with England and US banning nuclear weapons in space, Israel fought its Six-Day War against Syria, Jordan, Iraq, and Egypt, China became the world's

Corresponding author:
Ian Kerr, Canada Research Chair in Ethics, Law and Technology, University of Ottawa, Faculty of Law, Ottawa, ON K1N 6N5, Canada.
E-mail: iankerr@uottawa.ca

fourth thermonuclear power, and the first automated teller machine was installed in the United Kingdom—NBC aired "A Taste of Armageddon," the 24th episode in the first season of *Star Trek: The Original Series*.

The stardate is 3192.5. Captain Kirk and crew have beamed down to the planet Eminiar Seven.

KIRK: My mission is to establish diplomatic relations between your people and mine.
ANAN: That is impossible.
KIRK: Would you mind telling me why?
ANAN: Because of the war.
KIRK: You're still at war?
ANAN: We have been at war for five hundred years.
KIRK: You conceal it very well. Mister Spock?
SPOCK: Sir, we have completely scanned your planet. We find it highly advanced, prosperous in a material sense, comfortable for your people, and peaceful in the extreme. Yet you say you are at war. There is no evidence of this.
ANAN: Casualties among our civilian population total from one to three million dead each year from direct enemy attack. That is one reason, Captain, why we told you to stay away. As long as your ship is orbiting our planet, it is in severe danger.
SPOCK: With whom are you at war?
ANAN: The third planet in our system, called Vendikar. Originally settled by our people and now a ruthless enemy. Highly advanced technologically.
(*An alarm sounds*.)
ANAN: Please excuse me. Vendikar is attacking. Mea, care for our guests.
(*The Council adjourns into an adjacent room filled with computer equipment*.)

The Council adjourns to discuss the attack. Meanwhile, Captain Kirk and his crew are mystified. Their counterparts on the Starship Enterprise have been monitoring the planet from above and no signs of strife have been registered—no explosions, no radiation, no disturbances whatsoever. Yet the Eminiarian emissary assures them this is no game: one million deaths have just been registered. Before long, resident logician Spock begins to comprehend what has happened.

SPOCK: Computers, Captain. They fight their war with computers. Totally.
ANAN: Yes, of course.
KIRK: Computers don't kill a half million people.
ANAN: Deaths have been registered. Of course they have twenty four hours to report.
KIRK: To report?
ANAN: To our disintegration machines. You must understand, Captain, we have been at war for five hundred years. Under ordinary conditions, no civilisation could withstand that. But we have reached a solution.
SPOCK: Then the attack by Vendikar was theoretical.
ANAN: Oh, no, quite real. An attack is mathematically launched. I lost my wife in the last attack. Our civilisation lives. The people die, but our culture goes on.

KIRK: You mean to tell me your people just walk into a disintegration machine when they're told to?
ANAN: We have a high consciousness of duty, Captain.
SPOCK: There is a certain scientific logic about it.
ANAN: I'm glad you approve.
SPOCK: I do not approve. I understand.
…

Captain's log, delayed. The Enterprise, in orbit about Eminiar Seven, has been declared a casualty of an incredible war fought by computers. I, and my landing party, though apparently not included as casualties aboard the Enterprise, are confined on the planet's surface, awaiting what?[1]

When "A Taste of Armageddon" first aired in 1967, the prospect of "an incredible war fought by computers" was unquestionably the stuff of science fiction. And while the notion of computer-based decision-making may have been approaching reality, the idea of computer-fought war—virtual simulations dictating that millions of people must succumb to disintegration machines to keep the peace—required a suspension of disbelief. Such a ritualistic practice may have been tolerated by an ancient religious worldview,[2] but certainly not by our modern technological society. It was, therefore, not surprising when Captain Kirk, unwilling to accept Eminiar Seven's version of the Horatian fiction,[3] overpowered the guards and destroyed the war computers in what must have been perceived by the planet's inhabitants as a true Luddite uprising.

Forty-five years after Captain Kirk's glorious fictitious conquest,[4] our society faces the very real possibility of a similar kind of war—war fought by robots, not people. *A*

1. http://www.chakoteya.net/StarTrek/23.htm Star Trek: A Taste of Armageddon, unofficial transcript (NBC television broadcast, 21 July 1967).
2. Indeed, some Star Trek fans postulate that the historical Aztec "flower wars" inspired the episode: A Taste of Armageddon. http://en.wikipedia.org/wiki/A_Taste_of_Armageddon Wikipedia, the Free Encyclopedia. The Aztec flower wars did not involve battle in the usual sense. Instead, the Aztecs entered into an agreement with the Tlaxcalans to have ritual battles called xochiyaoyotl, the goal of which was not taking land, conquest, or death of the enemy, but simply to capture prisoners with whom each side could engage in ritual sacrifice. Not unlike the fictitious Eminiar and Vendikar civilizations, the Tlaxcalans prisoners volunteered themselves for ritual sacrifice: Hassig, *Aztec Warfare: Imperial Expansion and Political Control* (2005). In a Special Edition Extra for the Sci-Fi Channel, William Shatner suggests that the plot is also a metaphor for the sacrificial conscription of US soldiers in the controversial Vietnam war, which was in full swing at the time of broadcast: http://www.youtube.com/watch?v=CbmcVn5eHXU A Taste of Armageddon Star Trek Sci-Fi Channel Special Edition Extras, YouTube (April 5, 2010).
3. *Dulce et decorum est pro patria mori*—"It is sweet and fitting to die for one's country": Horace, Odes (III.2.13).
4. Voicing contrary opinions, there seem to be at least a few amongst the trekkies who think this episode was not one of the more glorious episodes that the original Star Trek had to offer: http://www.imdb.com/title/tt0708414/reviews Bogmeister, Not a Fan, IMDB. (17 July 2006); http://www.kethinov.com/startrekepisodes.php?id=172%20%E2%80%93%20this%20episode%20received%20a%204/10 Star Trek TOS- 1x23- A Taste of Armageddon; http://www.

Taste of Armageddon foretells a world that perceives war as inevitable, but its essential characteristics—e.g., physical disaster, loss of property, loss of knowledge and culture, disease, displacement, pain, anguish, starvation, and violent, bloody death—as evitable. It presents a certain "scientific logic": although we could never do away with war, we can use technology to minimize its impact on our day-to-day lives. Technology can make the conditions of war more tolerable. This, in turn, may result in more war. But it results in a much more bearable form of war. Through computers, war becomes civilized, more humane.

The basic rationale for today's robotic warfare is similar, though the looming robot wars will not be as neat and tidy as Star Trek's computer-fought war. In the virtual war between Eminiar Seven and Vendikar, aside from the inevitable prospect of random death by lottery, all of war's other atrocities become completely evitable. For both sides, civilian life during wartime carries on without disruption. Robotic warfare, on the other hand, does not enjoy symmetry in those characteristics of war that are evitable. The side without robots[5] will suffer physical disaster, loss of property, loss of knowledge and culture, disease, displacement, pain, anguish, starvation, and violent, bloody death. The side with robots will enjoy the spoils.

Semi-autonomous robots are already a well-entrenched component of military weaponry. And the project of developing lethal autonomous robots to replace human soldiers engaged in conflict is well under way, with an increasing annual global investment from the military and the private sector.[6]

Proponents[7] tell us that robotic warfare is merely an extension of human action in battlespace[8]—only with superior results and to nobler ends. The substitution of

avclub.com/articles/the-return-of-the-archons-a-taste-of-armageddon,25813/ Z. Handlen. "The Return of the Archons"/"A Taste of Armageddon", The A.V. Club, (27 Mar. 2009).

5. Or with fewer or weaker robots.
6. The investment and speed with which certain technologies are adopted is often jump-started by war. Singer cites a DARPA report that predicts the war on terrorism will do for military robotics what World War I did for Ford's Model T: in 1908, there were just 239 Model T cars sold but by 1918, there were over one million. Similarly, a defense consultancy called the Teal Group, which specializes in forecasting financial trends in war, expects worldwide spending on unmanned planes and computer-guided missiles to exceed $103.7 billion by 2016. Peter W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century*, eBook: (2009), p. 138.
7. Roboticist Ronald Arkin at Georgia Tech is a well-known advocate for the advantages of robot warfare: R. Arkin, "The Case for Ethical Autonomy in Unmanned Systems", *Journal of Military Ethics* 9 (2010), p. 332. John Canning makes similar arguments: J. Canning, "A Concept of Operations for Armed Autonomous Systems", Presentation to the 3rd Annual Disruptive Technology Conference, 6–7 Sept. 2010, at Washington, DC. The US military has employed ethicists and physicists including Patrick Lin and George Bekey to project just how advantageous advanced robotic armies could be: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697 P. Lin, G. Bekey & K. Abney, "Autonomous military robotics: risk, ethics, and design", Report for US Department of Navy, Office of Naval Research (2007).
8. In his book, *War X*, Tim Blackmore considers whether it is time for humans to leave what he terms "battlespace": T. Blackmore, *War X: Human Extensions in Battlespace* (Toronto: University of Toronto Press Inc., 2005).

automatons for humans in international conflict not only will save lives, we are told, robo-soldiers will outperform human soldiers physically, emotionally and ethically.[9] Robots are not vulnerable to the perils that plague humans on the battlefield: exhaustion, elevated emotions, or the need to seek retribution for the death of a comrade. Advanced sensory capabilities permit robots to cut through the fog of war—reducing confusion, friendly fire and other erroneous responses. Through a carefully programmed slave-morality, we can better ensure that robot warriors comport with international standards and the ethical rules of just war; unlike human beings, we can program robo-soldiers to be ethical. This, we are told, will reinforce and enhance international humanitarianism and reduce injustice in armed conflict. Like the computer-fought wars of science fiction, robot warfare promises to be civilized, more humane.

In this article, we examine the blurry lines between science and fiction in the policy discussions concerning the military use of lethal autonomous robots. We invoke two literary genres—science fiction and science and technology studies (STS). The focal point in the STS literature to date involves speculation about whether the next generation of military robots might in fact comport with the international laws of war. We contend that it is also important to consider whether and how the introduction of lethal autonomous robots might skew international humanitarian norms. We are less interested in the Asimovian question of whether robots can be programmed to adhere to the law than we are about how their participation in war might itself change the law by altering perceptions regarding the boundaries of permissible military force. In an attempt to better understand these things, we suggest that there is much to be learned from science fiction and STS.

In Part I, we survey the current state of the art in military robotics. Recognizing that today's semi-autonomous technologies require a "human in the loop," we suggest that it may not be long before human participation ceases to be a technological or military imperative. Part II considers the case for lethal autonomous robots and the technological project of programming ethical compliance with international humanitarian norms. In Part III, we demonstrate that none of these ideas are new; that their ethical, political and legal implications have been carefully explored for more than a half-century in science fiction, which has much wisdom to offer. Canvassing some of that literature, we articulate what we see as two predominant visions of robots. We further explore these particular visions of robots in Parts IV and V. More specifically, Part IV investigates the philosophical underpinnings of international humanitarian law's (IHL) fictitious approach to regulating the technologies of war: technological neutrality. We evaluate the problems inherent in that approach. In addition to its superficial and disingenuous treatment of the robotic technologies in question, we suggest that this approach permits a deterministic mode of thinking, expanding the scope of that which is perceived of as "necessary" once the technology is adopted. In Part V, we examine the implications of this in the IHL context, arguing that the "normative pull" of some emerging military technologies reshapes the rules regarding their permissible use. Consequently, we argue, even if lethal autonomous robots

9. R. Arkin, P. Ulam & A. Wagner, "Moral Decision Making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception", *Proceedings of the IEEE Special Issue on Interaction Dynamics at the Interface of Humans and Smart Machines*, 100 (2011), 571.

can be said to comport with IHL, they will operate as a force multiplier of military necessity, thus skewing the proportionality metric and amplifying new forms of destructive, lethal force. We conclude in Part VI by revisiting some of the insights offered by science fiction and by calling into question the appropriateness of IHL as the primary or exclusive means of regulating lethal autonomous military robots.

I. Military Robots

Robotic weapons can be broadly classified into three categories: (i) *remote-controlled weapons* (e.g. unmanned aerial vehicles that require a human operator's confirmation before the weapon's launch sequence can be engaged);[10] (ii) *semi-autonomous robots* (e.g. robot sentries that can accept sensory inputs and execute a specific action from a catalogue of set responses);[11] and (iii) *lethal autonomous robots* (e.g. the yet unrealized logical extension and oft-stated end goal of the current technology—machine systems capable of making tactical decisions and performing military operations independent of human input).[12]

Many of our most powerful weapon systems are already imbued with some degree of "autonomy" and, once initiated, are able to carry out operations independent of human control or oversight. For example, torpedoes[13] and cruise missiles[14]—precursors to today's Predator drones[15]—have long been able to determine (within a delimited range of options) the optimal speed and trajectory required to engage a target without human

10. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697 P. Lin, G. Bekey & K. Abney, "Autonomous military robotics: risk, ethics, and design", Report for US Department of Navy, Office of Naval Research (2007), p. 56.
11. Semi-autonomous robots have not departed completely from the standard of "human in the loop," but delegate increasing amounts of decision-making authority to the robot. Examples of this technology include modern iterations of drones, which operate independently and receive only destination coordinates from transmitters. Op. cit. p. 105.
12. R. Arkin, "Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture", Technical Report GIT-GVU-07-11 (2007), p. 4.
13. Torpedoes were one of the first technologies to self-select their targets, to some degree. American mathematician John Von Neumann and a team of engineers designed a torpedo that received feedback from its environment—much like a household thermostat—to home in on its destination: D'Amato, "International Law as an Autopoietic System," in R. Wolfrum and V. Roben (eds), *Developments of International Law in Treaty Making* (Berlin: Springer, 2005), p. 10.
14. Cruise missiles employ GPS technology to guide the missile to its target. The discovery of GPS was exalted more for its potential impact on weapons-guidance than for any other application. Modern cruise missiles combine GPS with additional feedback systems to confirm targets. Puckett, "In This Era of Smart Weapons, Is a State Under a Legal Obligation to Use Precision-Guided Technology in Armed Conflict?" *Emory International Law Review*, 18 (2004), 657.
15. Predator drones, a type of unmanned aerial vehicle, are one of the most well-known robotic weapons. About 27 feet in length, Predators resemble "baby planes" although they do not have cockpits. Predators are lightweight as they are constructed from composite materials, are less expensive than other military aircraft, and become a lethal stealth operator when equipped with Hellfire missiles. In their first year of operation by the American military, Predators fired missiles at 115 targets in Afghanistan. Singer, *Wired*, pp. 81–90.

presence and, in some instances, without human intervention.[16] A dramatic improvement over the fledgling application of air power in World War I (WWI), one example of such "smart" bombs are precision-guided munitions, which employ "laser, electro-optical, or infrared guiding systems that keep them on course towards their targets."[17] While precision-guided munitions have become increasingly predominant since first used by American forces in Vietnam, these smart bombs emerged into widespread usage and public consciousness after use by coalition forces in the 1990–1991 Persian Gulf War.[18] Their success at avoiding so-called "collateral damage," by preventing civilian casualties in targeted attacks in urban areas was highly publicized.[19] By the Gulf War, the circular error of probability of bombs dropped was a mere 10 feet, a dramatic improvement over the circular error of probability of 3,300 feet expected during WWI.[20] Smart weapons are now the standard for many military applications; indeed, some human rights advocates even argue that only sufficiently smart weapons should be permitted to attack within urban areas.[21]

A significant number of currently operational military robots fall within the category of unmanned aerial vehicles. By its own assessment, the US Department of Defense spent over $3 billion on these vehicles between 1990 and 2000, and another $4 billion from 2000 to 2010.[22] High profile Predator drone attacks on suspected terrorists in Pakistan, Yemen, and elsewhere have recently resulted in numerous front-page headlines.[23] In 2009, the US Air Force trained more remotely controlled aircraft pilots than actual fighter pilots.[24] New applications extend the already strong capabilities of the Predator. One such example, the Global Hawk, has been referred to as "the Predator's big brother."[25] It flies autonomously as opposed to being remotely piloted: an operator tells the unmanned aerial vehicle to take off with a mere click of a mouse. The Global Hawk

16. R. Sparrow, "Killer Robots", *Journal of Applied Philosophy*, 24 (2007), 64.
17. Infeld, "Precision-Guided Munitions Demonstrated Their Pinpoint Accuracy in Desert Storm; But Is A Country Obligated to Use Precision Technology to Minimize Collateral Civilian Injury and Damage", *George Washington Journal of International Law & Economics*, 26 (1992), 109.
18. Op. cit., 109.
19. Op. cit., 110.
20. http://www.ciaonet.org/olj/sa/sa_99bag02.html Bakshi, "Yugoslavia: Air Strikes Test of the Air War Doctrine", 22 *Strategic Analysis: A Monthly Journal of the IDSA* (1999).
21. Puckett, "Smart Weapons", 657.
22. Sparrow, "Killer Robots", 63.
23. See e.g., http://www.cnn.com/2012/03/30/world/asia/pakistan-drone-strike/index.html Kassim, "Drone strike hits Pakistani tribal region" CNN (30 Mar. 2012); http://www.guardian.co.uk/world/2012/apr/02/rise-of-the-drones-military-dilemma Norton-Taylor, "Rise of the drones poses growing dilemma for military: MoD confronts moral and legal issues as armed robots increasingly take warfare out of human control," *The Guardian* (2 Apr. 2012); http://www.reuters.com/article/2012/03/10/us-yemen-airstrike-idUSBRE82905V20120310 M. Ghobari & M. Mukhsaf, "Air strikes in Yemen kill 45 suspected Qaeda militants," Reuters, 10 Mar. 2012.
24. L. van Wifferen, "Alienation from the Battlefield: Ethical Consideration concerning Remote Controlled Military Robotics", MA thesis, Utrecht, The Netherlands: Universiteit Utrecht, 2011.
25. Singer, *Wired*, p. 92.

then carries out its pre-programmed mission by acquiring directions in real time from its onboard Global Positioning System (GPS) and operates independently until it returns and the pilot "hit[s] the land button."[26]

Land application of robotic weapons is more difficult and, as a result, less common. Robots can have difficulty navigating uneven terrain. Nonetheless, the US Army's Future Combat Systems project, which is currently in development, is aimed at developing a system for rapid deployment that would replace the current main battle tank with unmanned technology.[27] Early entrants to the realm of robotic land weapons include PackBot, a flagship product of the American robot company iRobot.[28] First used in rescue efforts on September 11, 2001 to adeptly navigate the carnage at Ground Zero, PackBot was later deployed to Afghanistan to act as a scout in treacherous cave systems.[29] Another land application, the Modular Advanced Armed Robotic System, is a combat robot manufactured by the QinetiQ Group that rolls on tank-like treads.[30] Its additional functionality includes day and night vision cameras, a four-barrel grenade launcher, and a 7.62 mm machine gun.[31] A semi-autonomous robot called the Samsung SGR-A1, a robotic sentry that can track multiple moving targets using infrared and visible light cameras, patrols the demilitarized zone between North Korea and South Korea.[32] This robot can identify a human target and fire upon it without human input, by detecting and identifying targets with a series of cameras, heat and motion sensors.[33] In its current deployment, a human makes the decision to fire the 5-millimetre light machine gun, but there is also an "automatic" mode where the robot can make its own decision.[34] South Korea aims to use this robot to shoot any human attempting to cross the demilitarized zone.[35]

There are a myriad of examples to showcase the advanced state of technology today. For instance, the Phalanx system for Aegis-class cruisers can "autonomously perform its

26. Op. cit.
27. Sparrow, "Killer Robots", 63.
28. iRobot purchased its name from the seminal Isaac Asimov novel. See generally: I. Asimov, *I, Robot* (New York: Bantam Dell, 1950). The company, founded by three MIT computer specialists, began its operations with small-scale government contracts and efforts at robotic toys. Its first major commercial success was the Roomba, the world's first mass-marketed robotic vacuum cleaner. Roomba was the progeny of a military robot, Fetch, which was designed by the US Air Force in 1997 to remove cluster bomblets from airfields. Singer, *Wired*, pp. 51–5.
29. Op. cit., p. 57.
30. http://online.wsj.com/article/SB10001424053111904070604576516591798551476.html Hodge, "Robots for Land, Sea, and Air Battles", The Wall Street Journal (19 Aug. 2011).
31. Op. cit.
32. http://robotzeitgeist.com/2006/11/samsung-techwins-sgr-a1-robot-sentry.html Samsung Techwin's SGR-A1 robot sentry video, Robotic Zeitgeist, Artificial Intelligence and Robotics Blog (Nov. 14, 2006).
33. http://andrewgibsondefence.wordpress.com/2010/10/17/should-killer-robots-be-banned-publisher-leftfootforward-org/ Gibson, "Should Killer Robots Be Banned?" LeftFootForward, (October 2010); see also: http://spectrum.ieee.org/robotics/military-robots/a-robotic-sentry-for-koreas-demilitarized-zone Kumagai, "A Robotic Sentry For Korea's Demilitarized Zone" IEEE (March 2007).
34. Op. cit.
35. Op. cit. See also: Arkin, "Governing Lethal Behavior", 5.

own search, detect, evaluation, track, engage and kill assessment functions."[36] Foster-Miller's has developed the Talon SWORDS (Special Weapons Observation Reconnaissance Detection System) platforms.[37] SWORDS vehicles autonomously move towards their targets using its GPS, while the firing of weapons remains the responsibility of a soldier that is located a safe distance away and remotely executes commands.[38] Lightweight miniature drones such as the RQ-11 Raven, with enhanced live-coverage capability, can be operated by remote or be programmed to operate autonomously.[39] Israel has stationary robotic gun sensor systems fitted with "fifty caliber machine guns and armored folding shields" positioned along its borders with Gaza in "automated kill zones."[40] While the current operation keeps a human in the loop to fire the weapon, autonomous operation is anticipated.[41] An especially notable development, the US Air Force's "Low Cost Autonomous Attack System," is specifically designed to "autonomously search for, detect, identify, attack and destroy theatre missile defence, surface to air missile systems, and interdiction/armour targets of military interest."[42] The system, which is equipped with radar and target recognition, can autonomously select between three different warhead configurations to ensure use of the best weapon for the job.

Systems such as these continue to improve: components become smaller, computer processing becomes more powerful and less expensive, and weapons capabilities become more and more adept.[43] The Moore's Law-ish[44] trajectory of military robotics is perhaps best evidenced by the American military operation in Iraq. When the forces first went into Iraq in 2003, only a handful of unmanned aerial vehicles were involved in the operation; that number is now 5,300.[45] At its outset, the operation had no robotic systems on the ground; now, there are over 12,000 such systems in place.[46]

---

36. G.E. Marchant et al., "International Governance of Autonomous Military Robots", *Columbia Science & Technology Law Review*, 12 (2011), 274.
37. Op. cit. See also: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697 P. Lin, G. Bekey & K. Abney, "Autonomous military robotics: risk, ethics, and design", Report for US Department of Navy, Office of Naval Research (2007), p. 12.
38. Marchant, "International Governance", 274.
39. Singer, *Wired*, p. 93. See also: http://www.army-technology.com/projects/rq11-raven/ 'RQ-11 "Raven Unmanned Aerial Vehicle, United States of America".
40. Arkin, "Governing Lethal Behavior", 5.
41. Op. cit.
42. Sparrow, "Killer Robots", 63.
43. Op. cit.
44. More than 40 years ago, Intel co-founder Gordon Moore forecasted the rapid pace of technology innovation. His prediction, popularly known as "Moore's Law," describes the doubling power of computer processing. Moore observed that transistor density on integrated circuits had doubled about every two years from 1957 to 1965 and predicted that they would continue to do so until at least 2020: See: Moore, "Cramming more components onto integrated circuits", *Electronics* 38, (April 19, 1965).
45. http://www.ted.com/talks/pw_singer_on_robots_of_war.html Singer, "Military Robots and the Future of War", Lecture delivered at TED (February 2009).
46. Op. cit.

Both technological and military standards regarding such weapons are currently premised on a "human in the loop"—for now, human beings are still the gatekeepers of military decision-making.[47] However, as the trajectory of many of the above-described technologies demonstrates, it may not be long before human participation ceases to be a technological or military imperative. More than 40 countries are in the process of developing autonomous weapons of various sorts.[48] Many of the systems under development will go far beyond "a 'fire and forget' system capable of determining its trajectory or pursuing its target to some limited extent."[49] We stand on the precipice of a military era that could require us to decide whether and to what extent we should delegate to machine systems the programming of missions, final targeting instructions, and even decisions about whether and when to pull the trigger or push the button.

II. The Case for Lethal Autonomous Robots

Lethal autonomous robots offer a seductive military advantage: their programmable nature means common human frailties could well be programmed out of next generation warriors. To some extent, robots in the field today outstrip their human counterparts in terms of sensing capabilities—not to mention recording and broadcasting. With highly developed sensor systems operating at incredible speeds of transmission, these machines can respond to information from many inputs simultaneously.[50] This enables real-time analytical surveillance of the battlefield, reducing the "fog of war."[51] Robotic senses are not clouded by human emotions such as fear, hysteria, anger and frustration.[52] They do not suffer from the human shortcoming of "scenario fulfillment," a propensity in human cognition to ignore or modify incoming information to mesh with their pre-existing beliefs and ideas—what Gary Marchant et al. describe as "a form of premature cognitive closure."[53]

---

47. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697 P. Lin, G. Bekey & K. Abney, "Autonomous military robotics: risk, ethics, and design", Report for US Department of Navy, Office of Naval Research (2007), p. 70.
48. U. Pagallo, "Robots of Just War: A Legal Perspective", *Philosophy & Technology*, 24 (2011), 315.
49. Sparrow, "Killer Robots", 64.
50. R. Arkin, "The Case for Ethical Autonomy in Unmanned Systems", *Journal of Military Ethics*, 9 (2010), 334.
51. Op. cit., 333. The fog of war refers to the uncertainty felt in a battlefield situation due to the volatile nature of that milieu and the unavailability of real time updates. This can render a military commander unaware of the force's—and the enemy's—current status and capabilities. Prussian military analyst Carl von Clausewitz is credited with having coined the term in his unfinished opus "Von Kriege" ("On War"). He writes: "The great uncertainty of all data in war is a peculiar difficulty, because all action must, to a certain extent, be planned in a mere twilight, which in addition not infrequently like the effect of a fog or moonshine gives to things exaggerated dimensions and unnatural appearance." See C. Von Clausewitz, *On War*, trans. and eds. M. Howard and P. Paret (Princeton, NJ: Princeton University Press, 1976).
52. Arkin, "Ethical Autonomy", 334.
53. Marchant, "International Governance", 276.

Today's robots are often employed in the types of situations wherein ordinary human weaknesses prove problematic. Since robots need not emulate the human tendency for self-sacrifice, they are better able to carry out operations conservatively in situations when the target has not necessarily been identified. Similarly, such robots can be used in a self-sacrificing manner without the moral reservation of a commanding officer or the need to steel itself against human survival instincts.[54] As Peter Singer notes (somewhat tongue in cheek) about the "death" of a PackBot—a key member of most American sentry teams operating in Iraq—when a robot dies, you do not have to write a letter to its mother.[55]

Yet beyond the advanced technological capabilities, it is the possibility of programming behavior that is often identified as a key advantage in autonomous systems. Ronald Arkin also suggests that robotic warriors could be made to perform more ethically than their human counterparts—that common human frailties could indeed be "designed out" of ethically programmed robots.[56] In wartime, human soldiers may disregard humanitarian requirements for a variety of reasons. In some instances, decisions may be emotionally charged, including the quest for revenge upon instances of high friendly losses or frustration at a poorly defined enemy.[57] There is also a recurring tendency to dehumanize the enemy, thereby rendering enemy soldiers apparently unworthy of humanitarian protections.[58] Other situations are more banal: when troops are immature or have been poorly trained and do not understand the laws of war. Such things can occur as a result of mundane circumstances such as a high turnover in the chain of command, weak leadership, or an issuance of unclear orders.[59] All of these shortcomings, Arkin and others argue, could be eliminated if the execution of war were to be delegated to the machines.[60]

In particular, Arkin contends that case-based reasoning, already employed in other intelligent robotic systems, could prove useful in this regard.[61] He believes that such military applications would eclipse the performance of soldiers on the battlefield, and he writes in a recent progress report: "I am convinced that they can perform more ethically than human soldiers are capable of."[62] According to Arkin, the fact that robots might sometimes fall short of the standard of the laws of war does not mean that they are worse than humans. Humans also fall short of this standard with what Arkin characterizes as "depressing regularity."[63] Using comprehensive system mapping and logical

54. Arkin, "Ethical Autonomy", 333. This would, of course, require a different programming approach than the one contemplated by Asimov's third law (set out in the text preceding note 89).
55. Singer, *Wired*, p. 52.
56. Arkin, "Moral Decision Making", 4.
57. Arkin, "Ethical Autonomy", 337.
58. Op. cit.
59. Op. cit.
60. Op. cit.
61. Arkin, "Governing Lethal Behavior", 12.
62. Op. cit., 7.
63. Arkin, "Ethical Autonomy", 1. And indeed, those who trumpet the value of human participation in war are confronted with statistics demonstrating that environmental factors lead

expressions, he describes the proposed functionality of lethal autonomous robots, including "architectural design options," that would inject moral reasoning capability into robots.[64]

Additionally, one of the strongest arguments in favor of lethal autonomous robots is the resulting decrease in human involvement in undesirable and perilous conflict. Simply put: why would we involve people in war if we don't have to? Fewer human soldiers mean fewer casualties—at least for one side of the battle.[65] Likewise, we are told, delegating these tasks to robots will eliminate much of the dangerous and undesirable work for military personnel—whose efforts can now be directed to higher order activities.

Of course, this 21st century rhetoric has a pedigree that must be unpacked in order to evaluate the case for lethal autonomous robots. Its origins lie in Athenian slave morality and its modern day progeny reside in contemporary science fiction.

III. The Robot Concept

When one considers the matter historically, the case for delegating undesirable or perilous tasks to robots is hardly a 21st century idea. It is well recognized as a popular, recurring meme in 20th century science fiction.

However, the roots of the robot concept go back to the Greek poets, whose ideas were subsequently formalized in 350 BCE, when Aristotle considered them in Book I, Part IV of his *Politics*. With all the playfulness of contemporary science fiction authors such as Jules Verne or Douglas Adams, Aristotle asked his readers to imagine a counterfactual world, one very different from his own:

> … if every instrument could accomplish its own work, obeying or anticipating the will of others, like the statues of Daedalus, or the tripods of Hephaestus, which, says the poet,
>
> "of their own accord entered the assembly of the Gods;" if, in like manner, the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves.[66]

Aristotelian robots—artificial slaves—eliminate the need for human labor. This insight would later ignite skyrockets in the minds of contemporary science fiction authors, and set in silicon his underlying vision of automation: the delegation of unpleasant or undesirable human activity to artefacts. However, Aristotle's robotic utopia was

soldiers to commit war crimes. For example, a report tendered by the Surgeon General of the 2006 American Operation Iraqi Freedom showed there was a tendency to disregard battlefield ethics training. Alarmingly, 10% of American soldiers that served in Iraq reported mistreating non-combatants and one-third of those surveyed would allow torture. See: Mental Health Advisory Team (MHAT), "IV Operation Iraqi Freedom 05-07", Office of the Surgeon Multinational Force-Iraq and Office of the Surgeon General United States Army Medical Command, 17 Nov. 2006, p. 36.

64. Arkin, "Ethical Autonomy", 1.
65. Singer, *Wired*, p. 86.
66. Op. cit.

relatively impoverished. Among other things, it lacked any assurances that artificial slaves capable of "anticipating the will of others" would not also be capable of insurrection or, at the very least, of carrying out harmful operations unintended by those who initially employed them. Contemplating these possibilities and safeguarding against them is the stuff of contemporary science fiction, from which two quite different visions of robots have been formulated.

As Ben-Naftali and Triger point out,[67] the first contemporary use of the word "robot" appears in a science fiction play written in 1920 by Karel Čapek, titled "R.U.R." (Rossum's Universal Robots). It is by now well-known that the term "robot" was in fact coined by Karel's brother, Joseph, as a derivative from the Czech word *robota*—used to describe a form of involuntary servitude in which low standing members of a social class were legally forced to work for the state or for aristocrats because they could not afford to pay taxes.[68] Setting the stage for an entire subgenre of science fiction devoted to robot insurrection,[69] Čapek imagines a society so deeply entrenched in capitalism that the laborers themselves become part of the means of production. Not merely in Marx's figurative sense, where the proletariat is exploited by the bourgeoisie, but quite literally—Rossum's factory fabricates[70] artificial people that are programmed to labor happily for the Rossum Corporation. However, the plan goes very badly awry once a hostile robot rebellion ensues.

Unquestionably, Aristotle's politics of servitude are inextricably linked to the perceived benefits of Rossum's robots. Consider a few lines from the play:

> *Domin:* The best sort of worker is the cheapest worker. The one that has the least needs. What young Rossum invented was a worker with the least needs possible.[71]
>
> …
>
> *Fabry:* So that they can work for us, Miss Glory. One robot can take the place of two and a half workers. The human body is very imperfect; one day it had to be replaced with a machine that would work better.

67. See: O. Ben-Naftali and Z. Triger, "Conditioning the Human Condition" in this symposium at p. 6.
68. Karel had originally thought about using the word "laboři" (connoting the idea of "laborer") but sought the advice of his brother, a painter, whose suggestion of "robota" better depicted the involuntary nature of the work: Čapek, *Lidove noviny*, 24 Dec. 1933, (Norma Comrada trans.) http://capek.misto.cz/english/robot.html; See also: Robertson, "Gendering Robots: Posthuman Traditionalism in Japan", in *Recreating Japanese Men*, S. Fruhstuck and A. Walthall (eds), (2011), p. 288.
69. See e.g.: Dick, "We Can Remember It for You Wholesale" in *The Magazine of Fantasy & Science Fiction*, E.L. Ferman (ed), (Apr. 1966)—the 1990 film "Total Recall" is based loosely on this work.
70. By a process we would today call "synthetic biology." The term was first used by Stéphane Leduc. See: Leduc, *Théorie physico-chimique de la vie et générations spontanées* (1910). For a contemporary description see: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2174633/ L. Serrano "Synthetic biology: promises and challenges", *Molecular Systems Biology*, 3 (2007), 158.
71. http://ebooks.adelaide.edu.au/c/capek/karel/rur/ Čapek, *R.U.R. (Rossum's Universal Robots)*, trans. by David Wyllie.

Busman: People cost too much.
Fabry: They were very unproductive. They weren't good enough for modern technology. And besides, … besides … this is wonderful progress that … I beg your pardon.
Helena: What?
Fabry: Please forgive me, but to give birth to a machine is wonderful progress. It's more convenient and it's quicker, and everything that's quicker means progress. Nature had no notion of the modern rate of work. From a technical point of view, the whole of childhood is quite pointless. Simply a waste of time.[72]

…

Domin: … now is the time when we can go back to the paradise where Adam was fed by the hand of God, when man was free and supreme; man will once more be free of labour and anguish, and his only task will once again be to make himself perfect, to become the lord of creation.[73]

…

Alquist: … The whole world, all the continents, all of mankind, all of it's [*sic*] just become one bestial orgy! No-one ever has to reach out his hand for food; he just stuffs it straight in his mouth without even needing to stand up. Haha, Domin's robots, they always take care of everything! And us human beings, the pinnacle of creation, we don't have to take care of work, we don't have to take care of children, we don't have to take care of the poor![74]

Using science fiction as a vehicle of enquiry, Čapek's R.U.R. raises interesting and important philosophical and ethical questions about Aristotle's utopian vision of delegating all unpleasant or undesirable labor to (intelligent) machines. This subject continues to develop today, not only within the science fiction genre but also in the field of contemporary moral philosophy.[75]

As moral philosopher, Stephen Petersen, describes it:

> Fiction involving robots almost universally plays on a deep tension between the fantasy of having intelligent robot servants to do our every bidding, and the guilt over the more or less explicit possibility that having such intelligent creatures do our dirty work would simply be a new form of slavery. The robot character is frequently a sympathetic antihero who gets mistreated by its callous, carbon-chauvinist human oppressors. Our guilt over such a scenario often manifests as a fear that the robots' servile misery would drive them to a violent and relentlessly successful uprising.[76]

72. Op. cit., act I.
73. Op. cit.
74. Op. cit.
75. See e.g.: Petersen, "Designing people to serve", in *Robot Ethics: The Ethical and Social Implications of Robotics*, P. Lin, K. Abney, and G. Bekey (eds), (2012) 283 [Petersen, "Designing"]; Petersen, "The ethics of robot servitude", 19(1) *Journal of Experimental and Theoretical Artificial Intelligence*, (2007), 43–54.
76. Petersen argues forcefully against the mainstream that there is in fact no moral problem with robot servitude, so long as the robots are hardwired to desire and prefer servitude: Op. cit., p. 283.

While there are many ways of telling stories about robots, and many variations even within these themes, we suggest this is the first of two relatively distinct visions of robots in science fiction. The catalogue of dystopic science fiction stories portraying a rebelling caste of *robota* is enormous and complex.[77] Those who take this narrative seriously, and many credible experts do, worry deeply about the project of developing lethal autonomous robots. Some believe that the future may not need us.[78] Of course, many others are not as concerned about an actual robot uprising.[79]

Unsurprisingly, there is another less destructive vision of robots that has developed in science fiction, which carries equal if not greater significance within the genre. This much more utopian vision of robots in science fiction imagines instead a peaceful, flourishing society in which robots work cooperatively alongside humans. This alterative narrative of "co-robotics"[80] was perhaps most famously promulgated by the science fiction of Isaac Asimov. In creating his robotic vision, Asimov was careful to respond to Čapek's vision of robots in R.U.R. and what Asimov called its "Frankenstein Complex"—"its gut fears that any artificial man they created would turn upon its creator."[81]

Asimov described what we are calling the second predominant vision of robots as follows:

> Beginning in 1939, I wrote a series of influential robot stories that self-consciously combated the "Frankenstein complex" and made of the robots the servants, friends, and allies of humanity.[82]

---

77. This, of course is the plot of *R.U.R*. Others in the genre include: Shelley, *Frankenstein; or, The Modern Prometheus* (2nd edn. 1823); *2001: A Space Odyssey* (Metro-Goldwyn-Mayer 1968); *Bladerunner* (Warner Bros. 1982); *Terminator* (Hemdale Film Corporation 1984); *Bicentennial Man* (Touchstone Pictures 1999); *Robocop* (Orion Pictures 1987); *Runaway* (Tri-Star Pictures 1984); *Transformers* (Paramount Pictures 2007); *Battlestar Galactica* (Original series: ABC: 1978; Re-imagined series: 2004–2007 NBC Universal Television, Universal Media Studios 2007–2009).
78. See e.g.: http://www.wired.com/wired/archive/8.04/joy.html Joy, "Why the future doesn't need us", *Wired*, 8 (April 2000).
79. The two of us included. Although we do find George Dyson's famous line quite compelling: "In the game of life and evolution there are three players at the table: human beings, nature, and machines. I am firmly on the side of nature. But nature, I suspect, is on the side of the machines." G. Dyson, *Darwin Among the Machines* (New York: Perseus Books, 1998), p. xi. For us, the reason why we ought not to treat robots as slaves has less to do with the unlikely prospect of robots becoming ethical beings capable of exacting revenge than it does with: (i) the human consequences that attach to treating humanoids as a slave class, and (ii) the unintended consequences of delegating decisions and relinquishing control to autonomous machines.
80. As it is styled these days: See, e.g.: http://robots.law.miami.edu/ian-kerr-and-jason-millar-on-delegation-relinquishment-and-responsibility-the-prospect-of-robot-experts/ J. Millar and I. Kerr, "Delegation, Relinquishment and Responsibility: The Prospect of Expert Robots" (2013).
81. Asimov, *The Bicentennial Man and Other Stories* (New York: Doubleday, 1976), p. 69.
82. http://www.e-reading.org.ua/chapter.php/81822/54/Azimov_-_Robot_Visions.html. Asimov, "The Machine and the Robot", in *Robot Visions* (New York: Penguin Books, 1990).

Perhaps an even more apt depiction of Asimov's robotic vision is expressed through the voice of one of his most interesting characters in the Robot Series, Dr. Susan Calvin. Learning that her interlocutor is only 32 years old, she proclaims:

> Then you don't remember a world without robots. There was a time when humanity faced the universe alone and without a friend. Now he has creatures to help him; stronger creatures than himself, more faithful, more useful, and absolutely devoted to him. Mankind is no longer alone. Have you ever thought of it that way?
>
> …
>
> To you, a robot is a robot. Gears and metal; electricity and positrons. Mind and iron! Human-made! If necessary, human-destroyed! But you haven't worked with them, so you don't know them. They're a cleaner, better breed than we are.[83]

Asimov's idea was that universally coding the machines to comport with certain social rules could mitigate the Frankenstein complex. Taking Aristotle to the next level, the elite could not only use robots to their great advantage but the general public could also be made to trust robots. Programming all robots with an obedient, slave morality would allay people's fears, ensuring that robots are "more faithful, more useful, and absolutely devoted." Not to mention, ubiquitous.

This approach was codified in what became famously known as Asimov's "Three Laws of Robotics," requiring that:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.[84]

Numerous commentators (sometimes including Asimov himself) have, from time to time, rather carelessly treated the three laws as though they were simply human laws to be followed by human-like beings. However, Asimov's three laws are better understood as a system of automated permissions; software code that prescribes and proscribes by way of architecture what people can and cannot do with robots. Understood in this sense, the actual prose of the three laws is really just an abstract ethical code for roboticists to guide their programming. As such, the three laws are a precursor to an extremely

83. Asimov, *I, Robot*, p. xi.
84. http://www.ebooktrove.com/Asimov,%20Isaac/Asimov,%20Isaac%20-%20Robot%2006%20-%20Robots%20&%20Empire.pdf?9fd8a6e3e5d474fad5f9001399292f58=783ce58e014bf2de6ee912a401a904b6 Asimov subsequently posited the need for a "zeroth law": "a robot may not harm humanity, or, by inaction, allow humanity to come to harm." Asimov, *Robots and Empire* (New York: HarperCollins, 1985), part V, c.18.

powerful set of ideas made famous right around the bong of the new millennium by cyberlaw proponent Lawrence Lessig—that "code is law."[85]

The basic strategy goes like this. Instead of exclusively developing ethical or legal norms of human conduct and imposing them on people in difficult situations (e.g., as expressed in IHL) have programmers install the rules directly into the operating systems of the machines as well, so that people simply cannot use the robots to carry out prohibited conduct. For example, in accordance with Asimov's "first law," if a human tried to enlist a robot to injure another person in a manner contrary to the first law, the robot would refuse, shut down, or otherwise render itself incapable of carrying out the command. It was never easy to convince or trick an Asimovian robot into wrongdoing. Asimov's robots' positronic brains were hardwired to do no evil.

Interestingly, the "three laws" approach of Asimov's science fiction is very quickly becoming science fact in a growing field called roboethics.[86] This is part of the approach adopted by Arkin et al., outlined above in Section II. Applying roboethics to a real-life military example, a robot sentry in an international conflict might be programmed to automatically disable when it identifies a civilian as the target of a mission. This is necessary in order to comply with IHL, which requires weapon systems to be able to discriminate between combatants and non-combatants.[87] Arkin and other roboethicists believe that ethical programming of lethal autonomous robots in this way not only will assuage the Frankenstein Complex but, also, that by replacing human fighters with ethically programmed machines, we can in fact achieve civilized, more humane outcomes.

Of course, this was precisely the underlying vision of Asimov's three laws. As Dr. Susan Calvin retorted in the following exchange with Stephen Byerley:[88]

Byerley: You're the U.S. Robot's psychologist, aren't you?
Calvin: Robopsychologist, please.
Byerley: Oh, are robots so different from men, mentally?
Calvin: Worlds different. … Robots are essentially decent.[89]

85. Reidenberg, "Lex Informatica", *Texas Law Review*, 76 (1998), 553; Lessig, "Code and Other Laws of Cyberspace and Code 2.0" (2006).
86. Created by Gianmarco Veruggio in 2002 roboethics is a field of study concerned with the behavior of humans, how humans design, construct, use and treat robots and other artificially intelligent beings. See: http://www.roboethics.org/index_file/Roboethics%20Roadmap%20 Rel.1.2.pdf G. Veruggio, "Roboethics Roadmap 1.2", European Robotics Research Network (January 2007); G.Veruggio, "Views and visions in Robotics", Hearing at the Italian Senate's 7th Permanent Commission, (Rome 2002); J. Searle, "Minds, brains and programs", *Behavioral and Brain Sciences*, 3 (1980), 417; A. Winfield, "Five roboethical principles – for humans", *New Scientist*, 210 (2011), 32.
87. For an interesting application to the case of lethal autonomous robots, see: http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA534697 P. Lin, G. Bekey & K. Abney, "Autonomous military robotics: risk, ethics, and design", Report for US Department of Navy, Office of Naval Research (2007), p. 50.
88. A politician accused of being a robot.
89. Asimov, *I, Robot*, p. 178.

By which she meant: programmed by humans so that they could not act otherwise.

Reifying Asimov's science fiction, subsequent generations of roboticists actually believe that a careful integration of ethically programmed robots into combat would lead to civilized, more humane outcomes. Are Arkin et al. correct in thinking wars that involve more robots and fewer humans are better likely to comport with the ideals that undergird the international legal framework for just war? Alternatively, will an increasing automation and robotization of warfare be used, unjustifiably, to redefine what counts as "humane" according to international humanitarian norms?

To answer these questions, we begin the next part by briefly describing the relevant principles of IHL. Although the correct application of these principles is fundamental to lawful outcomes in the actual use of lethal autonomous robots, in our view, it is even more essential to understand IHL's strategic framework for the regulation of emerging military technologies—the doctrine of technological neutrality. Part IV seeks to carefully unpack this doctrine in order to demonstrate its limitations.

IV. IHL and the Doctrine of Technological Neutrality

The history of humanitarian law is unquestionably tied to the development of new technology: as war technologies become increasingly advanced and capable of greater destruction, laws are put in place to limit that destruction.[90] These laws have been conceived in two distinct streams: laws of general application that apply to all instances of warfare to ensure the imperative of humanity modulates how war is waged; and specific rules/laws that prohibit or limit the use of particular weapons, e.g. chemical and biological weapons. Within this framework, three key humanitarian principles are employed: (i) proportionality, (ii) distinction, and (iii) military necessity.

Proportionality requires that any collateral injury to civilians and private property during a military operation must be balanced against the military advantage to be gained by carrying out that operation.[91] Distinction addresses the weapon's capacity to distinguish between combatants and non-combatants. It requires that weapons and operational strategy select targets to minimize collateral damage.[92] The final core principle, military necessity, is best explained by its own name: the attack must be militarily necessary, i.e., the state must only defend itself and seek to guarantee its future security in executing the military operation—nothing more.[93] There are associated responsibilities to produce

---

90. http://www.icrc.org/eng/resources/documents/statement/new-weapon-technologies-statement-2011-09-08.htm Kellenburger, "IHL and New Weapon Technologies". Thirty-fourth Round Table on Current Issues of IHL, Keynote Address (September 8, 2011).
91. E. Cannizzaro, "Contextualizing Proportionality: jus ad bellum and jus in bello in the Lebanese War", *International Review of the Red Cross*, 88 (2006), 785.
92. Y. Dinstein, *War, Aggression and Self-Defence* (3d) (New York: Cambridge University Press, 2001), p. 119.
93. C. Greenwood, "Historical Development and Legal Basis" in D. Fleck, ed, *The Handbook of Humanitarian Law in Armed Conflicts* (New York: Oxford University Press, 1995), p. 36.

intelligence on the effects of the weapons used, the number of civilians that could be affected, and whether they could take cover before an attack.[94]

These well-known principles, understood in light of a broader principle of humanity, are meant to guide all tactical decision-making by a state engaged in an armed conflict. Adherence to these principles is designed to both minimize destruction during wartime and facilitate the peace building process after a ceasefire is reached: the hope is that reconciliation is made easier without malicious or indiscriminate attacks during the conflict. The critical legal question is usually framed in the same manner, irrespective of which weapon is selected: does the specific use comply with the above-mentioned principles?[95]

But this is not the only question.

We contend that it is even more important to consider how the introduction of lethal autonomous robots into battlespace might impact international humanitarian norms. Might the participation of lethal autonomous robots in a war alter the norms of IHL?

To better comprehend this possibility, we offer an analytic suggestion. IHL can be understood as adopting a particular strategic framework for the regulation of emerging military technologies. This approach is known in other disciplines as the doctrine of *technological neutrality*.[96] Rather than implementing sector specific rules or laws that are tailored to the functionality or capabilities of particular technologies, IHL rests on a set of foundational principles.[97] These principles are said to be "neutral" with respect to any given technology. On this approach, military technologies are not regulated categorically or by class but through a determination of whether a particular implementation or use conflicts with the underlying principles of IHL.[98] Consequently, a military

94. UNHRC, Report of the Special Rapporteur on extrajudicial, summary, or arbitrary executions, 14th Sess Supp No 40, UN Doc A/HRC/14/24/Add.6 (2010), para. 29; M. Schmitt, "Military Necessity and Humanity in International Humanitarian Law: Preserving the Delicate Balance", *Virginia Journal of International Law*, 50 (2010), 828.
95. UN Doc A/HRC/14/24/. Report of the Special Rapporteur on extrajudicial, summary, or arbitrary executions (20 May 2010), para. 79.
96. Technological neutrality is a commitment to the idea that laws should be framed generally, as opposed to being designed for a specific technology. This view is intrinsically linked to a value-free understanding of technology itself: "technology is deemed 'neutral' without valuative content of its own." See: http://www.sfu.ca/~andrewf/books/critical_theory_of_technology.pdf Feenberg, "Critical Theory of Technology" (1991), p. 5. Moreover, technology-neutral language permits a one-size-fits-all solution to regulatory problems. From a regulatory perspective, employing generic regulations will prove more efficient as it will avoid modifying its entire policy position when the passage of time inevitably delivers new technologies. This ensures not only a non-discriminatory approach that treats different technologies equally when creating regulations themselves, but also a sustainable approach to lawmaking that will prevent frequent revisions or laws based around particular technologies. van der Haar, "Technological Neutrality; What Does It Entail?" Mar. 2007 TILEC Discussion Paper No. 2007-009, pp. 22–3. See generally: www.lex-electronica.org/docs/articles_236.pdf Ali, "Technological Neutrality", *Lex Electronica*, 12 (2009).
97. For example, the principles of distinction, proportionality, military necessity, humanity, etc.
98. As discussed, there are several international frameworks that prohibit or limit the use of specific weapons. These include: Declaration Renouncing the Use, in Time of War, of Explosive

technology will only be limited or restricted if the manner in which it must be used or the results that it achieves is in conflict with international humanitarian principles.

The underlying approach has been successfully adopted in the regulation of other technologies. In the context of electronic commerce, technological neutrality

> refers to statutory tests or guidelines that do not depend upon a specific development or state of technology, but rather are based on core principles that can be adapted to changing technologies. Since technological change is constant, standards created with specific technologies in mind are likely to become outdated as the technology changes.[99]

Consequently, the standards adopted are deemed to be technology-neutral. In other words, the same standards can then be applied across a range of technologies. From a regulatory perspective, employing generic regulations will prove more efficient as it will avoid modifying its entire policy position when the passage of time inevitably delivers new technologies.[100] This ensures not only a non-discriminatory approach that treats different technologies equally when creating regulations themselves, but also a sustainable

Projectiles Under 400 Grammes Weight, Saint Petersburg, (11 Dec. 1868 or 29 Nov. by the Julian Calendar) UKPP 659 (LXIV); Protocol for the Prohibition of the Use of Asphyxiating, Poisonous or Other Gases and Bacteriological Methods of Warfare, 17 June 1925, UKTS 24; Convention on Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques, 10 Dec. 1976, UNTS 151; Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects (and its Protocols: Protocol I: non-detectable fragments, Protocol II: mines, booby-traps and other devices, Protocol III: incendiary weapons) 10 Oct. 1980, 1342 UNTS 137. An additional Protocol (IV) on blinding laser weapons was enacted on 13 Oct. 1995; Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction, 3 Sept. 1992 1974 UNTS 45; Convention on the Prohibition of the Use, Stockpiling, Production, and Transfer of Anti-Personnel Mines and on Their Destruction, 18 Sept. 1997 2056 UNTS 211; Convention on Cluster Munitions, 30 May 2008, CCM/77. Still, these conventions remain the exception: in most situations, there is insufficient political will to generate such agreements, even in the most outwardly self-evident cases, e.g. nuclear weapons. In the post-bellum period, global powers were unable to reach an accord on the legality of nuclear weapons, despite (or perhaps because of) the havoc they wreaked in Hiroshima and Nagasaki. The specter of the Cold War and prospective nuclear arsenals loomed large. Accordingly, conversations about outlawing weapons are tied to fears that those weapons will be used against you once you agree to get rid of your own.

99. http://law.berkeley.edu/journals/btlj/articles/voI16/geist/geist.pdf Geist, "Is There A There There? Towards Greater Certainty for Internet Jurisdiction", *Berkeley Tech LJ*, 16 (2002), 1359. See also: http://www.wipo.int/mdocsarchives/WIPO_EC_CONF_99/WIPO_EC_CONF_99_SPK_24%20C_E.pdf Herrmann, Secretary of United Nations Commission International Trade Law (UNCITRAL), "Establishing a Legal Framework for Electronic Commerce: the Work of the United Nations Commission on International Trade Law", conference paper, presented to the International Conference on Electronic Commerce and Intellectual Property (WIPO), September 14–16, 1999, p. 45.

100. van der Haar, "Neutrality", 22–3.

approach to law making that will prevent frequent revisions or laws based around particular technologies.[101]

In an effort to avoid reinventing the legislative wheel for each and every emerging technology, the doctrine of technological neutrality, as a policy objective, tells us that we ought to guide our laws not by the technological state of the art but on the basis of sound legal judgments about the underlying functions that the various relevant technologies aim to achieve. While policymakers in electronic commerce law have enjoyed success in relying on this doctrine to date,[102] it is instructive to consider why technological neutrality might not be well suited for at least some emerging technologies.

Consider an example from the field of data protection (which also currently relies on technological neutrality as a guiding principle). Rather than making a separate privacy law for video rental records, another for surveillance cameras, yet another for facial recognition systems, and still another for social network sites, etc., the global approach to data protection has been to develop a core set of functional principles that are meant to adapt to various technologies which will emerge over time. Consequently, most domestic privacy laws are derivative of a set of eight such principles, first articulated by the Organization for Economic Cooperation and Development in the 1980s.[103] Focusing on fair information practices like collection limitation, purpose specification, use limitation and the like, the flexibility of privacy law's technology-neutral approach has been quite remarkable in its ability to regulate data collection, use, and disclosure across an array of technologies not previously anticipated when the data protection principles were themselves first enunciated.

Despite its success to date, technological neutrality is no panacea. More recently, disruptive technologies have begun to emerge that undermine or otherwise turn-on-its-head one (or more) of data protection's core regulatory principles. For example, *ubicomp*—a set of sensor networks and automating technologies devised to eliminate the need for human interaction during a series of information transactions[104]—disrupts the general data protection requirement of "knowledge and consent" in the collection, use or

101. Op. cit., 22.
102. See e.g.: http://www.uncitral.org/enindex.htm UNCITRAL Model Law on Electronic Commerce With Guide to Enactment (1996), and its Canadian counterpart, the Uniform Electronic Commerce Act: http://www.law.ualberta.ca/alri/ulc/current/euecafin.htm.
103. The ten principles of fair protection appended in Schedule I of Canada's Personal Information Protection and Electronic Documents Act (PIPEDA) are modeled after the privacy principles developed by the Organization for Economic Cooperation and Development (OECD). Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5. See also: http://oecdprivacy.org/ OECD Privacy Principles. In the United States, similar principles originated in a 1973 government report as a system for protecting personal privacy in report keeping systems. See: http://bobgellman.com/rg-docs/rg-FIPShistory.pdf Robert Gellman, "Fair Information Practices: A Basic History", Version 1.92, June 24, 2013.
104. Ubiquitous computing, known colloquially amongst technophiles as "ubicomp", refers to a future where digital devices are integrated so seamlessly into our daily existence that no one notices their presence. Mark Weiser, widely considered to be the father of ubiquitous computing, begins his seminal piece "The Computer for the 21st Century" with this vivid description: "The most profound technologies are those that disappear. They weave

disclosure of personal information. The practical impossibility of obtaining regular and meaningful consent for every *ubicomp* transaction may result in the need for some new sector specific regulatory approaches, should our society fully embrace the model of ubiquitous computing. As we suggest below, a similar need may arise in IHL in the case of lethal autonomous robots. Put colloquially, some technologies change the game.

The basic assumption[105] of the doctrine of technological neutrality is further problematized when it is reduced to the more general proposition that technologies are themselves neutral. This form of technological neutrality treats technological tools as value-free empty vessels, ready to fit the uses of their users.[106] "Guns don't kill people," goes the maxim. "People kill people."

Although not frequently consulted in the literature on military technology and IHL, STS has had much to say about this second form of technological neutrality. As Neil Postman so eloquently put it,

> Embedded in every technology there is a powerful idea, sometimes two or three powerful ideas. Like language itself, a technology predisposes us to favor and value certain perspectives and accomplishments and to subordinate others. Every technology has a philosophy, which is given expression in how the technology makes people use their minds, … in how it codifies the world, in which of our senses it amplifies, in which of our emotional and intellectual tendencies it disregards.[107]

Despite the very compelling work of Langdon Winner and so many others in the field who have sought to demonstrate that artefacts have politics,[108] we continue to "disregard the fact that many technologies determine their own use, their own effects, and even the kind of people who control them. We have not yet learned to think of technology as having ideology built into its very form."[109] To retell one of the best and most famous examples from the literature, let us consider in some detail the automatic mechanical tomato harvester, a remarkable device developed by agricultural researchers at the University of California in the late 1940s.

Most people would ask: could something as straightforward as a mechanical tomato combine really have an ideology built into its very form?

themselves into the fabric of everyday life until they are indistinguishable from it." See: M. Weiser, "The Computer for the 21st Century", *Scientific American*, 265 (1991), 94. Ubicomp technologies will be foundational to the coming era of robotics.

105. Namely, that we can, for particular legal purposes, treat all or even most technologies the same.
106. Feenberg, "Critical Theory of Technology", 5.
107. N. Postman, *The End of Education: Redefining the Value of School* (New York: Vintage Books, 1996), pp. 192–3.
108. L. Winner, "Do Artifacts Have Politics?" in L. Winner (ed), *The Whale and the Reactor: A Search for Limits in an Age of High Technology* (Chicago, IL: The University of Chicago Press, 1986), p. 19.
109. J. Mander, *Four Arguments For The Elimination of Television* (New York: HarperCollins, 1978), p. 350.

The mechanical tomato combine cuts the plant from the ground, shakes the fruits loose, and—in newer models—sorts the fruits by size. In order to accommodate the shaking, scientists have bred new varieties of hardier, sturdier, and less tasty tomatoes. The machine replaced the handpicking system wherein human harvesters went through the fields multiple times to pick ripe tomatoes and leaving unripe fruits for later harvest. The machines reduce the cost of tomatoes by $5–7 per pound. The cost of these machines (~$50,000) means that only highly concentrated forms of tomato farming can afford this type of harvesting. Consequently, with the advent of the combine, the number of tomato farmers dropped from ~4,000 in the early 1960s to ~600 in 1973. An estimated 32,000 jobs were lost as a direct result of mechanization.[110]

Winner contends that the broad adoption of the mechanical tomato harvester ultimately shifted the norms of tomato farming in California and, indeed, the nature of the tomato itself. His observations offer interesting insights that can be applied to our considerations about the future military adoption of lethal autonomous robots, and its implications for the norms of IHL and the nature of war itself. But, before delving further into that, let us first take a more careful look at Winner's analysis by imagining a corollary set of norms that would seek to regulate farming (and its environmental impact) by similar means.

We can imagine an environmental farming law that employs principles similar in nature to IHL. Eco-just farming might entail that those engaged in farming must likewise adhere, *mutatis mutandis*, to principles of: "proportionality," "distinction," and (let's call it) "agricultural necessity."

The principle of distinction therefore requires that acts of farming should be directed only to agricultural products and not non-agricultural vegetation subsisting in its natural environment.

We can imagine entrenching these corollary norms as a means of safeguarding the environment against the potential evils of modern agriculture. And yet it is not difficult to see that the outcomes described by Winner would be reproduced. If we can assume that the means by which the harvester is employed does not implicate other non-agricultural vegetation subsisting in its natural environment (Winner did not discuss this), the adoption and use of the mechanical tomato harvester would easily be justified in terms of its agricultural necessity and said proportionality.

The agricultural necessity principle, in this context, requires that an agricultural intervention must be intended to help achieve an agricultural objective and the resulting ecological harm must be proportional and not excessive in relation to the concrete and direct agricultural advantage achieved through the intervention. In this case, the agricultural need to feed many people in and outside of California is met through the enormous increase in yield, the significant reduction in cost, and the incredible overall efficiency in tomato production. Meeting agricultural objectives in this way would be understood as a social benefit that eclipses any sacrifices to the marginal practice of rural agricultural culture. Applying the Doctrine of Double Effect,[111] since the foreseeable harm to rural agricultural practice was inextricably intertwined with the agricultural objectives of

110. Winner, "Artifacts", 22.
111. T. Aquinas, *Summa Theologica* (II-II, Qu 64, Art .7).

increased productivity, efficiency, and cost-effectiveness in the harvesting process, and since the introduction of the mechanical tomato harvester was primarily intended to achieve these overall goods, its adoption will be justified.

Proportionality, in this context, requires that any ecological harm during an agricultural operation must be balanced against the anticipated agricultural advantage gained by carrying out that operation. Recall that proportionality requires that an agricultural intervention seeking to fulfill an agricultural objective must not be undertaken if the ecological harm is known to outweigh the anticipated agricultural advantage. As is often the case, the harms cannot be measured until the technology is in place. Once it is in place, we are no longer talking about some neutral cost-benefit analysis. The very adoption of this cumbersome and expensive machine system, the layoff of crew-upon-crew of farm workers, the reconfiguration of farmland, etc., are themselves an assertion of a political preference for one set of values (productivity, efficiency, cost-effectiveness) over another (homegrown, authentic, domestic). The introduction of the mechanical tomato harvester is an assertion of one way of life and the negation of another.

This is perhaps most evident in the ultimate effect of the tomato harvester. Here, we are not merely referring to the displacement of a set of rural agricultural practices or the reshaping of social relations in tomato production but, in fact, to an entire way of being in the world. We are referring to the transformation of the tomato itself.

Through a sleepy haze, we almost need to remind ourselves: what was the original agricultural objective in introducing the mechanical tomato harvester?

Upon wakeful reflection, the objective was to *better harvest* tomatoes, which, it turns out, is a completely different goal than to harvest *better tomatoes*. The only shortcoming of this otherwise incredible agricultural device was that it could not handle soft, juicy (delicious) tomatoes. The introduction of this technology necessitated plant breeders to come up with "new varieties of tomatoes that are hardier, sturdier, and less tasty than those previously grown."[112]

Asleep at the switch, we practically forget that by permitting this technology we also permit it to determine its own use.

> Shielded by the conviction that technology is neutral and tool-like, a whole new order is built – piecemeal, step by step, with the parts and pieces linked together in novel ways – without the slightest public awareness or opportunity to dispute the character of the changes underway. It is somnambulism (rather than determinism) that characterizes technological politics … Silence is its distinctive mode of speech.[113]

It is because of this propensity towards technological somnambulism that, during and after the proportionality analysis, we tend to be dozy in recognizing that the entire balancing act was essentially dictated by the characterization of the objectives and the perceived "necessities" of the situation. "We *need* a new form of tomato," the engineer later tells us. Of course, this was not originally an agricultural necessity—but it became

112. Winner, "Artifacts", 22.
113. Winner, *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought* (Cambridge, MA: The MIT Press, 1978), p. 324.

perceived as such just as soon as the social investment in mechanical tomato harvesters was made. Technology can shape our perceived needs. So much the worse for tasty tomatoes.

It is in this sense that technology is not neutral and can be used (if we let it) to reshape social norms. As Winner was so clear to point out, this observation does not entail technological determinism.[114] Rather, it recognizes *that what technology makes possible has the power to generate in our minds what may later be perceived of as necessary*. But only as a result of the adopted technology, which subsequently permits (but does not necessitate) the cart to drive the horse. How else would it seem necessary to produce a kind of tomato that never before existed? The so-called agricultural necessity for a new tomato is not solely the consequence of adopting the mechanical tomato harvester. The ability to see the new tomato as an agricultural necessity is only perfected when plant breeders actually invent a technique for creating one.

At the same time, technology can also induce a related form of dogmatic-slumber[115] that permits another crucial fallacy to occur—an illogical inversion of Kant's famous insight that "ought implies can."[116] Through a strange form of grammatical alchemy we mistakenly come to believe that because the technology makes something possible, it

---

114. See generally: Heilbroner, "Do Machines Make History?", *Technology and Culture*, 2 (1961), 335; Shaw; "The Handmill Gives You the Feudal Lord: Marx's Technological Determinism", *History and Theory*, 18 (1979), 155; Hansen, "The Technological Interpretation of History", *Quarterly Journal of Economics*, 36 (1921), 72. Winner articulated his view as follows:

> It may seem that the view I am suggesting is that of technological determinism: the idea that technological innovation is the basic cause of changes in society and that human beings have little choice other than to sit back and watch this ineluctable process unfold. But the concept of determinism is much too strong, far too sweeping in its implications to provide an adequate theory. It does little justice to the genuine choices that arise, both in principle and practice, in the course of technical and social transformation. Being saddled with it is like attempting to describe all instances of sexual intercourse based only on the concept of rape. A more revealing notion, in my view, is that of technological somnambulism. For the interesting puzzle in our times is that we so willingly sleepwalk through the process of reconstituting the conditions of human existence.

L. Winner, "Technologies as Forms of Life", in D. Kaplan (ed), *Readings in the Philosophy of Technology* (New York: Rowman & Littlefield, 2003), p. 107.

115. In the preface to *Prolegomena to Any Future Metaphysics*, Kant famously "thanked" David Hume for his arguments about causality, which Kant credited as being "the very thing which many years ago first interrupted my dogmatic slumber and gave my investigations in the field of speculative philosophy quite a new direction." The implausibility of Hume's arguments on cause and effect led Kant to the realization that there were other ways of thinking of things a priori. He rejected Hume's reluctant conclusion that reason is deluded about cause and effect and set out to use Hume's argument as "a spark from which light may be obtained" by catching "some inflammable substance." I. Kant, trans. and ed. by G.Hatfield, *Prolegomena to Any Future Metaphysics: With Selections from the Critique of Pure Reason* (Cambridge: Cambridge University Press, 1997), p. 10.

116. The maxim "ought implies can" is a form of transcendental idealism that leaves open the possibility that we have free will. This ethical framework mandates that an individual must be logically able to perform any action that he or she is morally obligated to perform. As

also makes it necessary. *Our perceived needs are thereby (re)shaped by our sense of what is possible*. This propensity is crucial to remember in the application of balancing provisions. *How technology shapes our perceptions will have a significant effect on our understanding of what is proportional and the means by which we apply principles like agricultural necessity*.

Applying Winner's example to the military context will further assist in demonstrating the limitations of the doctrine of technological neutrality. Returning to where we began, Star Trek's computer-fought war provides a compelling illustration.

In keeping with the general approach of IHL, computer simulated (virtual) warfare will be understood as neutral in two senses. First, virtual war is not currently subject to specific international agreements that prohibit or limit its use and will therefore be treated exactly the same as any other technology subject to the laws of general application—despite its horrific implications.[117] Second, assuming that the virtual war machines are properly programmed to discriminate between combatants and non-combatants and that the algorithm for determining who is killed by way of collateral damage is unbiased (if not random), it might plausibly be claimed that virtual war machines are value-neutral in the sense that they are disinterested in the particular outcomes of computerized battle.

The military necessity requirement would be regarded as fulfilled on the basis that virtual battle is understood to be the only way to prevent the phenomenal devastation of actual battle. Citizens who perceive virtual war as a military necessity will claim that those identified by the computers as casualties have a moral and legal duty to report to the disintegration machines.

The proportionality calculus would generate a similar outcome. Once the technology is in place, the painless death of a fraction of the population, fairly chosen, would be regarded as outweighing the physical disaster, loss of property, loss of knowledge and culture, disease, displacement, pain, anguish, starvation, etc. that would otherwise be suffered by two entire civilizations engaged in an all-out interplanetary war.

Like the ironic effect of the tomato harvester on the tomato itself, the politics of computerized battle is perhaps most evident in its transformation of the very nature of war. One must recall that the original military objective of computer-fought war was to remove human beings from the realm of harm. And yet we see the outcome in fact

---

Kant asserts: "For the moral law commands that we ought to be better human beings now, it inescapably follows that we must be capable of being better human beings." Moreover: "The action to which the 'ought' applies must indeed be possible under natural conditions." Kant, *Religion Within the Boundaries of Mere Reason* 6:50. See also Kant, *Critique of Pure Reason* A548/B576, 473. In the cyberspace context, Lawrence Lessig cautions against the use of sweeping rhetoric about the nature of technology, since it can lead to deterministic conclusions. He cites the frequent mention of the very "essence" of the web being a place that cannot be regulated, that its "nature" is to resist regulation. Yet, if there is any place where nature should have no rule, it is in cyberspace: a place that has been entirely constructed. Lessig characterizes this as "the fallacy of 'is-ism' – the mistake of confusing how something is with how it must be." L. Lessig, *Code: Version 2.0* (New York: Basic Books, 2006), p. 31.

117. And, let's assume for the purposes of this analysis that it will not become subject to such an agreement.

increases the ambit of harm immensely by creating a culture of fear in a society plagued by a persistent state of war, constantly subject to the threat of imminent, needless death. To Captain Kirk, and anyone else in a position to see beyond the perceived necessity of the circumstances, the death-by-disintegration of millions each year is tragically evitable. One can almost imagine Kirk appropriating Langdon Winner's great line: "It is somnambulism (rather than determinism) that characterizes [Eminiar Seven's] technological politics."[118] Wake up!

"But, we *need* to build disintegration machines," one imagines the architects of virtual war retorting. Of course, disintegration machines were not originally a military necessity. But they became perceived as such just as soon as the social investment in virtual warfare was made. The Star Trek example reaffirms that what technology makes possible has the power to generate in our minds what may later be perceived of as necessary. After all, it is extremely difficult to imagine the citizens of Eminiar Seven and Vendikar embracing the duty-to-be-disintegrated if they hadn't already viewed computer-fought war as a military necessity. This example reminds us, once again, that how technology shapes our perceptions will have a significant effect on our understanding of what is proportional and the means by which we apply principles like military necessity.

In the following part, we shift our focus back to the interaction between lethal autonomous robots and the norms of IHL. First, we consider in general the normative pull of military technologies, arguing that the mere introduction of some military technologies can actually alter prior norms and practices. Second, we attempt to show how shifting battle norms might occur through the introduction of lethal autonomous robots. We argue that robotic technologies act as a force multiplier in the determination of military necessity, thus amplifying the amount of permissible destructive force in carrying out an operation.

V. Lethal Autonomous Robots and the Norms of IHL

1 The normative pull of military technologies

While roboethics and the prospect of programming lethal autonomous robots to comport with international law are fascinating and fast emerging fields of study, an important but less well understood line of enquiry examines how international humanitarian norms are influenced and implicated through the adoption and use of military technologies.

Like the advent of the mechanical tomato harvester in rural Californian agricultural communities discussed above in Part IV, the introduction of a new military technology can reshape norms within military culture. Consider, for example, shifting standards in submarine warfare.[119] Humanitarian ideals have long informed the norms of

118. Winner, "Autonomous", 324.
119. London Procès-Verbal Relating to the Rules of Submarine Warfare Set Forth in Part IV of the Treaty of London of 22 April 1930, 6 Nov. 1936, UKTS 29 (1936), Cmd. 5302 (Eng. Fr.). Peter Asaro also uses this example to exemplify problems associated with lethal autonomous robots. See P. Asaro, "How Just Could A Robot War Be?" in P. Brey, A. Briggle and K. Waelbers (eds), *Current Issues in Computing and Philosophy* (IOS Press, online, 2008),

naval warfare between surface vessels. Conflicts on the high seas were accompanied by standard responsibilities that aimed to preserve humanity in any altercations.[120] For example, military vessels were prohibited from attacking merchant ships and were instead required to capture and escort merchants to port. Likewise, after a naval battle, the successful vessels were required to rescue survivors by bringing them aboard.[121] The 1930 London Naval Treaty and its 1936 successor codified the law in this area.[122]

The historical example of the submarine illustrates the vulnerability of technological neutrality: as norms evolve and nations point to state practice to justify actions that stray away from—or are even in direct contravention of—international agreements, such conventions run the risk of becoming "blue law."[123] Customary international law may purport to fill this vacuum; yet since international law depends on the development of norms based on the behavior of sovereign nation states, law can be transformed by a collective omission or new practice. And this can be achieved through the introduction of a new technology that "forces" new practices.

The 1936 London Naval Protocol reaffirmed that submarines had the same duties as surface vessels.[124] Consequently, submarines had a responsibility to comport with the longstanding obligations imposed on ships.[125] Attempting to follow these rules was not only impractical for submarines but also had the effect of imposing near impossible responsibilities. Lacking comparable crews to surface vessels, accompanying a merchant ship to port was not something submarines could feasibly do. Moreover, during World War II, German submarines were relegated to great depths for both their own safety and strategy since the Allied Forces controlled the surface.[126] Even if a submarine were to surface after a battle, the space constraints in a small cabin scarcely large enough for the

p. 59. However, this example originates in Michael Walzer's seminal work on just war. As a corollary to the "sink on sight" unlimited submarine warfare practices employed by German forces during World War II (WWII), German submarines stopped following the duty set out in the 1936 London Protocol to provide for the safety of the survivors of a sunken ship. This was justified under the auspices of military necessity: submarines are exposed to great danger if they have to surface and fulfill the obligations of surface vessels. The argument was made that the only alternative was to not use submarines at all or to use them ineffectively, which would have given control of the sea to the British navy. See Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations* (New York: Basic Books, 1977), p. 147.

120. Op. cit.
121. J. Gilliland, "Submarines and Targets: Suggestions for New Codified Rules of Submarine Warfare", *Georgetown Law Journal*, 73 (1984), 981.
122. Op. cit., 978.
123. Jane Gilliland describes a blue law as being one which is both clear on its terms and be clearly violated by an accused, but is unenforced because of changed conditions and long-term disregard by the community. Gilliland, "Submarines", 989.
124. London Procès-Verbal Relating to the Rules of Submarine Warfare Set Forth in Part IV of the Treaty of London of 22 April 1930, 6 Nov. 1936, UKTS 29 (1936), Cmd. 5302 (Eng. Fr.).
125. Gilliland, "Submarines", 989.
126. Op. cit., 981.

existing crew and machinery meant that taking on two or three more people was out of the question.[127]

The most fundamental incompatibility between submarine operation and the constraints imposed on surface vessels went to the core of naval strategy: submarines were intended for stealth. They were used in situations where stealth was paramount. Surfacing would negate the defined military objective for which they were deployed. Before long, in the tumultuous context of World War II, any assumption that submarines would comport with the Second London Naval Treaty disappeared. Accordingly, the treaty fell into disuse. New norms around the behavior of submarines emerged that were based on the way submarines were already being used in warfare. These norms were predicated—at least in part—on the way the technology had been designed.

Although the submarine did not unilaterally undermine the humanitarian requirements of the high seas, the very design of submarine technology did determine its own use with respect to those rules. Though such an outcome may well have been serendipitous (it is not like submarines were purposefully built to get around these rules), this example illustrates that international humanitarian norms can indeed be circumvented by design. In at least some instances, a state could avoid humanitarian obligations by implementing a technology that is said to fulfill an important military objective whose very operations would be undermined by complying with the norm.

Jane Gilliland cautions, "the law of armed conflict for submarines subrogates military necessity to humanitarian goals, and in so doing threatens the achievement of the humanitarian goals it seeks to protect."[128] The submarine example showcases how the advent of new technology may sculpt international norms into what is easily practicable when the technology employed becomes the new widespread practice. Since customary international law is formed in part by state practice, this is troubling. The codifications that later occurred took into account the practices that had already unfolded on the battlefield. Failing to account for technological change thereby weakens the staying power of the codification of international law.

Of course, one could also imagine a very different and much more humane historical outcome for the submarine. The fact that submarine vessels cannot easily rescue overboard combatants (who likely came to be that way because of the torpedo fired at them by the submarine) might just as easily be understood as a reason against their deployment rather than a reason in favor of excusing submarines from otherwise enforceable humanitarian obligations. In part, what history tells us is that the case in favor of military necessity is a strong one and that the technologies said to be necessary in carrying out important military objectives are not easily interfered with.

2 Lethal autonomous robots as a force multiplier of military necessity

Both the submarine and the mechanical tomato harvester were game changers in terms of what was subsequently seen as necessary and proportional in naval battle and tomato

127. Op. cit.
128. Op. cit., 991.

agriculture. The basic mechanism in each case was previously explained in Part IV above: when a disruptive technology changes the nature of what is possible, it also expands the scope of inclusion for what can prospectively be perceived of as "necessary."[129] The consequences of this for international law are significant. If norms can be shifted in a manner mandated by the technology, then its potential to transform international law—*where practice becomes principle*—is enormous.

The power to induce a shift in norms in this way has led many academics to register concern about technology's influence over international law. Colin Picker expresses his worry as follows:

> Perhaps most problematic is the [fact] that technology is a determinate force that acts as an invisible hand creating, shaping and destroying international law. Failure to handle such a powerful force will result in policy makers essentially abdicating the international regime to technology.[130]

With many references peppered throughout his article to "the invisible hand of technology,"[131] Picker seems at first blush to be suggesting that the appropriate underlying philosophical worldview for international law is technological determinism—the idea that technology determines social outcomes.[132] Picker later explicitly denies this, claiming: "I am not arguing, however, in favor of technological determinism. Technological

129. Better prediction technology makes possible an argument that pre-emption is necessary. For instance, the much maligned "Bush doctrine" justifies preemptive self-defense in the context of the "War on Terror" through advanced technological prediction capability that better informs government agencies about the threat of a terrorist attack. In his first public speech on the topic at a graduation ceremony at West Point, Bush stated: "If we wait for threats to fully materialize, we will have waited too long. … We must take the battle to the enemy, disrupt his plans, and confront the worst threats before they emerge[;] … our security will require all Americans to be forward-looking and resolute, to be ready for preemptive action when necessary to defend our liberty and to defend our lives." See: http://georgewbush-whitehouse.archives.gov/news/releases/2002/06/20020601-3.html "President Bush Delivers Graduation Speech at West Point" (June 1, 2002). See also: http://www.state.gov/documents/organization/63562.pdf "The White House, The National Security Strategy of the United States of America", (2002), pp. 14–15.

130. Picker, "A View from 40,000 Feet: International Law and the Invisible Hand of Technology", *Cardozo Law Review*, 23 (2001), 151.

131. As Colin Picker explains: "The fact that modern communications technology is forcing policy makers to take global constituencies into account is a reflection of a perception that technology is an irresistible force. This is not a new insight. Indeed, thirty years ago, L.F.E. Goldie argued that 'the sciences are all-pervasive in international law … and have a direct condition, if not determinative, effect both upon existing rules and upon the progressive development of new rules.' This insight is analogous to Adam Smith's metaphor of the role of the invisible hand of the market on a country's economy. In the technology and international law context, the metaphor would observe that technology operates as an invisible hand on international law, guiding and shaping its development." Op. cit., 201.

132. The phrase "technological determinism" was allegedly coined by economic industrialist Thorstein Veblen. See: http://socserv2.mcmaster.ca/~econ/ugcm/3ll3/veblen/Engineers.pdf

determinism implies a stronger and more comprehensive relationship between technology and international law than I would assert exists. Policy makers [*sic*] can ignore technology, but at a tremendous cost."[133]

Our position is somewhat different. To us, it is not as though the invisible hand of technology magically removes all other social outcomes or possibilities, or that "it" somehow punishes those who do not respond accordingly with heavy costs. Rejecting Picker's strange metonym, we contend that certain social uses of technology can reconstitute our perceptions so that we will not easily experience other existing possibilities.[134]

It is perhaps trite to say that international law was intentionally constructed to provide an extremely flexible framework.[135] Picker's more interesting claim is that a primary reason for doing so is to accommodate the protean nature of technology. Like the more recent technology-neutral frameworks used in electronic commerce and data protection law (both of which are derived domestically from international models), the core design of IHL is consistent with promoting, rather than restricting innovation. On this approach, we do not restrict innovation, only particular uses of it. While the technology in question is itself conceived of as neutral, the framework said to regulate it is not: it is designed to encourage and accommodate the overall use of technology.[136]

When value-neutral approaches are applied to deeply value-laden technologies, the results can be disingenuous. The four *Geneva Conventions* underlying IHL, which codified existing customary international law, were concluded in 1949. Animating this process was not only the aftermath of WWII, but the specter of weaponization in the newly arrived nuclear age. The process was concluded on August 12, 1949, less than a month after the then-USSR test detonated its first nuclear weapon.[137] Yet, efforts to categorically prohibit such weapons of mass destruction were met with resistance. Unlike

---

T. Veblen, *Engineers and the Price System* (2001), p. 38. For a fulsome survey of philosophers with varying view on technological determinism, see Bimber, "Karl Marx and the Three Faces of Technological Determinism", *Social Studies of Science*, 20 (1990), 333.

133. Picker, "View," 203.
134. N. Postman, *Technopoly: The Surrender of Culture to Technology* (New York: Vintage, 1993).
135. The multiple sources of international law in the ICJ Statute, which include treaty law, customary law, general principles of law recognized by civilized nations, and judicial decisions and writings of highly qualified publicists demonstrates that international law is not intended as a single behavioral code, but instead as a flexible entity that recognizes a pluralistic legal perspective. See Statute of the International Court of Justice (1945), 59 Stat. 1055, Art. 38(1). Moreover, the creation of assumedly concrete and binding treaty law is accompanied by planned review conferences to determine whether the scheme is sufficient and whether changes are necessary. For example, though the Rome Statute came into force in 2002, a review conference was mandated for seven years later to propose and discuss amendments to flexibly respond to its implementation. See: http://www.kampala.icc-cpi.info/ International Criminal Court, "The Review Conference for the Rome Statute".
136. For a similar argument see, S. Lavi, "Cloning International Law: The Science and Science Fiction of Human Cloning and Stem-Cell Patenting" in this symposium.
137. http://www.icrc.org/eng/resources/documents/misc/5krbdw.htm Gordon-Bates, "The ICRC and the nuclear weapon: the story of an uncomfortable paradox" (March 18, 2003).

previous consensus over combat gases and biological weapons, the international community was unable to establish the political will to prohibit the development of nuclear arsenals.[138]

It was not surprising—in the golden age of technological neutrality[139] – to see that very approach adopted by the International Court of Justice (ICJ) in its *Advisory Opinion on the Legality of Nuclear Weapons*.[140] The ICJ was resistant to the idea that one technology could be expressly forbidden and instead clung to the precepts of international law: the weapon itself was not illegal; the acts that could be committed with that weapon were illegal.[141] And provided the weapon was used in a manner that satisfied the legal test for proportionality, distinction, and military necessity, there was no need to outlaw it outright.

Despite the inherent flexibility of IHL, it is still difficult to imagine circumstances in which the use of a nuclear weapon could satisfy humanitarian norms. Nuclear weapons will, in their present form, be consistently unable to discriminate between civilians and combatants. Accordingly, any use of a nuclear weapon that satisfies the criterion of distinction would be an operation calculated to annihilate an entire area. As we have seen, such an act is clearly inhumane and almost certainly disproportionate to any act it purportedly responds to.[142] Moreover, and perhaps even more troubling, allowing a nuclear weapon to remain within the *arsenal of possibility* might permit its use to be justified prospectively under the guise of military necessity in subsequent situations. As Judge Higgins describes in the *Nuclear Weapons* decision, questions of numbers of casualties or inflicted suffering "must be resolved as part of the balancing equation between the necessities of war and the requirements of humanity."[143]

With all of this, we see that the framework for balancing international humanitarian norms is sufficiently malleable to permit destruction and lethal force. While it is true that

138. Op. cit.
139. The mid-1990s also saw international action towards the creation of technology neutral e-commerce legislation. Though there was an identified need to regulate electronic commerce, the international community remained reticent to the development of specific standards due to the propensity of the technology to change rapidly and the possibility of legislative obsolescence before the provisions even came into force. Moreover, an international agreement could only function if countries the world over, with different technologies in practice, would agree to a convention. The United Nations Model Law on Electronic Commerce therefore adopts a minimalistic, technology-neutral approach. See Official Records of the General Assembly, Fortieth Session, Supplement No. 17 (A/40/17)(1996). Concurrently, global data protection regimes also began to gain similar momentum, with the prevailing attitude being that policies needed to be technology neutral to withstand the test of time. The European data privacy directive, Directive 95/46/EC, was adopted in 1995. See: http://www.edps.europa.eu/EDPSWEB/edps/EDPS/Dataprotection/Legislation "Legislation", European Data Protection Supervisor.
140. Legality of the Threat or Use of Nuclear Weapons, ICJ Rep 226 (1996).
141. Op. cit., paras. 74–87.
142. Conceivably, one might argue that it would be proportionate to use a nuclear weapon in response to another nuclear attack.
143. Legality of the Threat or Use of Nuclear Weapons, ICJ Rep 226 (1996), paras. 14–20.

any military action must be constrained within the parameters of a proportional response, the overall potential for destruction is unquestionably augmented by the existence of certain destructive and lethal technologies with advanced capabilities.

As an illustration, let us imagine these norm-conflicts along a continuum. At one end of the continuum are outcomes premised solely on humane or humanitarian grounds (the principle of humanity). At the other end are outcomes that focus exclusively on carrying out destructive or lethal military objectives (the principle of military necessity). Somewhere in the middle, where these two norms are in direct conflict, the adoption of a new technology is often sought as a military solution.[144]

But the introduction of such a technology is often (to use a military metaphor) a "force multiplier"[145] of military necessity. As we have suggested, the ability to capitalize on military possibilities created by such technologies raises the stakes in terms of possible military objectives and operations that were unfathomable prior to the emergence of the technology. This allows us to "sleepwalk" towards a perspective that sees various uses of these technologies as a military necessary for resolving present and future armed conflict, even if they result in more death and destruction.

There is an interesting connection here between the concept of military necessity and the worldview of technological determinism that we reject. If one listens closely to the justifications often given in support of the military necessity in the use of a particular technology, there is a false strand of determinism running through it. The ICJ commits this very fallacy in contemplation of a situation wherein a State's "very survival would be at stake," anticipating that nuclear weapons may be the only recourse in "an extreme circumstance of self-defence."[146] It is this deterministic thread built into the very fabric of military necessity that makes it a force multiplier.

How might this show itself in the case of lethal autonomous robots?

In its most utopian form, military robotics seek to remove humans from battlespace, anticipating fewer people fighting and fewer casualties. In almost all conceivable situations, however, casualties would remain—it is only friendly casualties that are reduced by replacing one side's soldiers with robots. Failing to acknowledge this reality risks callousness in attack. Either way, the existence of lethal autonomous robots will surely be a force multiplier of military necessity in terms of the general military objective of reducing friendly casualties. "If we have expendable mechanical mercenaries that we can send into battle instead of our children, how could we *not* do so?"

Of course, once we do so, the use of lethal autonomous robots will have a profound effect of lowering the threshold for entry into war: war will be (domestically) perceived

144. The case for lethal autonomous robots set out above in Section II provides an excellent example of this.
145. The term "force multiplier" refers to a factor that significantly enhances the effectiveness or strategic advantage of a particular force. In a military context, technology can result in force multiplication that achieves an objective that would traditionally have required a much larger force. Arkin expects lethal autonomous robots would result in force multiplication by having robots and soldiers side-by-side on the battlefield. See Arkin, "Governing Lethal Behavior", 13.
146. Nuclear Weapons, para. 96.

of as easier both politically and logistically if there are fewer people involved. Fewer soldiers need to be recruited; fewer deaths have to be justified to win public support of the war effort, fewer injuries, less property damage, etc. As Captain Kirk astutely observed in the case of virtual war:

> Death, destruction, disease, horror. That's what war is all about, Anan. That's what makes it a thing to be avoided. You've made it neat and painless. So neat and painless, you've had no reason to stop it.[147]

As Kirk identifies, computer-fought war has a double effect: it impacts the traditional *jus in bello* rules for a fair fight on the battlefield while the capacity of the technology simultaneously influences a nation's political decision to go to war. The same can be said for a robot-fought war.[148] Again, one can view this as a force multiplier of military necessity. "If we have robots on the ground that can carry out an important military operation with few or no friendly casualties, why *wouldn't* we engage the enemy on this mission?"

Since lethal autonomous robots will still interact with human military collaborators, the force multiplier of remote controlled warfare can also have individual psychological effects on the soldiers engaged in it, increasing sympathy for military objectives and military necessity by minimizing our empathy regarding circumstances no longer seen through the lens of humanitarian ideals.[149] Today's warriors often fight from some cubicle, operating aircraft remotely, and developing a "Playstation mentality" with regard to the waging of war.[150] Recent YouTube footage of Predator drone attacks set to music and shared and celebrated amongst soldiers online, further distances the acts from the actors.[151] Increasing both the physical and psychological distance between soldiers and their targets not only dampens respect for human life—it also makes it easier to follow military objectives, especially those perceived of as necessary. After all, those being killed are only ever encountered as pixels on a screen.

This asymmetric element of robotic warfare is not only dangerous; it also conceptually challenges the foundations of war by skewing the balance between humanitarian ideals and military necessity. Doesn't the very nature of conflict change if one side's soldiers are never actually in danger? It is already the case that an American soldier can serve an entire tour of duty in Afghanistan or Iraq, work 16-hour days, and still eat lunch every day at Carl's Jr. just outside the gates of his or her Nevada Air Force base.[152] While it may be appealing to

147. http://www.chakoteya.net/StarTrek/23.htm Star Trek: A Taste of Armageddon, unofficial transcript (NBC television broadcast, 21 July 1967).
148. See e.g., Pagallo, "Just War", 303.
149. See e.g., Singer, *Wired*; Asaro, "Robot War"; Sparrow, "Killer Robots"; van Wifferen, "Alienation".
150. van Wifferen, "Alienation", 38.
151. These videoclips, many of which are freely available on YouTube, are colloquially known as "war porn." Peter Singer cites a particularly egregious example of a clip of catastrophic explosions being set to the song "I Just Wanna Fly" by Sugar Ray. http://www.ted.com/talks/pw_singer_on_robots_of_war.html Singer, "Military Robots and the Future of War," Lecture delivered at TED (February 2009).
152. Singer, *Wired*, p. 85.

imagine being at war with an enemy without experiencing casualties, it also solicits a larger question: if one side has no people in harm's way, is it truly a war?[153] And, more to the point for present purposes, without a significant level of human investment, will that side be able to see, understand, and inculcate the humanitarian norms in tension during battle? The advent of lethal autonomous robots has the potential to greatly exacerbate these risks in a manner that could be profoundly destabilizing to the framework of armed conflict.

VI. Evitable Conflicts, Inevitable Technologies?

We have suggested that it is crucial to recognize the philosophical underpinnings and implications of IHL's purportedly technology-neutral approach. Though this framework treats each technology under consideration as though it is neutral, we have engaged science fiction and STS literature to demonstrate that the current framework uncritically and unquestioningly encourages and accommodates the development and use of emerging technologies. We believe that a failure to recognize and unpack the values embedded into the design of the framework itself, let alone those embedded in the robotic technologies under consideration, can lead to a mistaken and deterministic mode of thinking that fallaciously treats unjustifiable, lethal operations as though they are a military necessity. We have offered a possible explanation for how this might occur: when a disruptive technology changes the nature of what is possible, it also expands the scope of inclusion for what can prospectively be perceived of as "necessary" in carrying out (military) objectives.

Asimov, whose predominant vision of robotics provides the blueprint for first generation lethal autonomous military robots, was astutely aware of the means by which robotic technology can reshape our perceptions of what is necessary. He wrote about this many times within the Robot Series, most intelligently in a short story titled, "The Evitable Conflict."[154] And, yet, his staggering insights have fallen off the radar of those engaged in today's policy debates about the implications of using lethal autonomous robots in battle.

Written in 1950, "The Evitable Conflict" imagines a global economy that relies (literally and figuratively) on a *deus ex machina* solution to the problems generated by a robot society. As World Coordinator, Stephen Byerley described it:

> The ending of every other problem had merely given birth to another. Our new worldwide robot economy may develop its own problems, and for that reason we have the Machines. The Earth's economy is stable, and will remain stable, because it is based upon the decisions of calculating machines that have the good of humanity at heart through the overwhelming force of the First Law of Robotics.[155]

In a policy discourse virtually identical to today's real life discussions about delegating military operations to lethal autonomous robots, the society Asimov asks us to imagine

153. http://www.economist.com/blogs/babbage/2010/10/robots_war "Robots at war: Drones and democracy", The Economist (October 1, 2010).

154. Asimov, "The Evitable Conflict", in *The Complete Robot* (New York: HarperCollins/Voyager, 1983), p. 450. Despite correcting them over and over, students almost always call this story "The Inevitable Conflict", which I find most telling.

155. Op. cit., p. 450.

has recognized that machine performance outstrips human thinking and doing on a number of fronts. In an effort to avoid *perceived* inevitable conflicts, such as war (famine, poverty, unemployment, pollution, destruction, etc.), key global economic and political decision-making is delegated to highly capable, utilitarian "Machines" designed to maximize human happiness. As a result, the world's population has come to witness no unemployment, no overproduction, no shortages, etc. Waste and famine are now merely words in history books. "And so the question of ownership of the means of production becomes obsolescent. Whoever owned them (if such a phrase has meaning), a man, a group, a nation, or all mankind, they could be utilized only as the Machines directed. Not because men were forced to but because it was the wisest course and men knew it."[156] Humankind has placed its faith in the Machines.

As Murphy's Law[157] would *only seem* to have it, the plot subsequently depicts various high level (human) personnel discovering that something has very badly gone awry, that the Machines have been "off" in their calculations. Of course, this should not be possible since the Machines' positronic brains are hardwired to follow the three laws. Consequently, World Coordinator, Stephen Byerley, is charged with investigating what has gone wrong. He consults each of the four Regional Coordinators and then Dr. Susan Calvin, robopsychologist at US Robots and Mechanical Men Inc.

After rejecting various hypotheses proposed by the Regional Coordinators regarding Machine-failure, Byerley discusses with Calvin the possibility that the Machines have been tampered with by the "Society for Humanity"—a small but powerful group, "who feel themselves strong enough to decide for themselves what is best for them, and not just to be told [by robots] what is best for others."[158]

After entertaining Byerley's suspicions, Calvin assures him that the Machines are several steps ahead of the game and that the Machines must have purposefully allowed the Society's plots, creating just enough turmoil so that the Society might destroy itself.

> … It is not the "Society for Humanity" which is shaking the boat so that the Machines may be destroyed. You have been looking at the reverse of the picture. Say rather that the Machine is shaking the boat – *very* slightly – just enough to shake loose those few which cling to the side for purposes the Machines consider harmful to Humanity.[159]

Thus, despite the protections supposedly afforded by the three laws, Byerley and Calvin come to recognize the unintended consequences of delegating key human decision-making to the Machines. In spite of the first law or—rather, *not* in spite of it—the Machines purposely skewed certain economic results, knowing that the humans, perceiving error, would attempt to "override" machine-based decisions. In other words, the Machines were intentionally producing "perceived errors" as a preemptive set of corrective measures for their correctly predicted human errors.

156. Op. cit.
157. Typically stated as, "Anything that can go wrong will go wrong." Though, as Bill Joy has noted: "Actually, this is Finagle's law, which in itself shows that Finagle was right." See Joy, "Future".
158. Asimov, "Evitable Conflict", 466. These are Asimov's luddites.
159. Op. cit., 468.

Given that even the cleverest attempts to overthrow the Machines only result in more data for the Machines to consider, the Machines ultimately twist the first law in order to assume total control over all humanity—this being the Machines' interpretation of "not allowing human beings to come to harm."

> "But you are telling me, Susan, that the 'Society for Humanity' is right; and that Mankind *has* lost its own say in its future."
>
> "It never had any, really. It was always at the mercy of economic and sociological forces it did not understand – at the whims of climate, and the fortunes of war. Now the Machines understand them; and no one can stop them, since the Machines will deal with them as they are dealing with the Society, – having, as they do, the greatest of weapons at their disposal, the absolute control of our economy."
>
> "How horrible!"
>
> "Perhaps how wonderful! Think, that for all time, all conflicts are finally evitable. Only the Machines, from now on, are inevitable!"[160]

Asimov imagines that delegation of decision-making to the Machines can in fact render war and human strife "evitable." However, he suggests that transforming the deep nature of such conflicts comes at a tremendous cost—rendering human dependency on the Machines "inevitable."

Hence, for Asimov, the inevitable becomes evitable, and the evitable, inevitable.

Such perceptual transformations, we posit, are not limited to science fiction.

Acknowledgements

Ian Kerr: I would like to thank the incredible Orna Ben-Naftali both for the opportunity to participate in this special symposium on Science Fiction and International Law and for the extremely valuable suggestions on an earlier draft of our article. Thanks also to my wonderful colleagues Jane Bailey, Chloe Georas, David Matheson, Madelaine Saginur, Ellen Zweibel, Jason Millar, and the super six pack: Eliot Che, Hannah Draper, Charlotte Freeman-Shaw, Sinzi Gutiu, Katie Szilagyi and Kristen Thomasen for their inspiring thoughts on this important emerging subject. Penelope Simons also offered extremely insightful and highly useful feedback on a previous draft; many thanks to them. Finally, I deeply appreciate Charlotte Freeman-Shaw, Beatrice Bozinovski, Zac Delong, Andrew Bigioni, Stephanie Low and Laura Freitag for maintaining grace under pressure and for the very high quality of their research.

Katie Szilagyi: My sincere thanks goes to Ian Kerr for this amazing opportunity, for his mentorship, and for his constant encouragement to go confidently in the direction of my dreams. I echo the above thanks to our wonderful "Robotnik" colleagues for their collective wisdom and thoughtful comments. Special thanks also to William Wong for his excellent editorial eye and to Jon Khan for always helping me to maintain my gameface.

160. Op. cit., 469.

Funding

The author wishes to extend his gratitude to the Social Sciences and Humanities Research Council and the Canada Research Chairs program for the generous contributions to the funding of the research project from which this article derives.