ORIGINAL ARTICLE

INTERNATIONAL LEGAL THEORY

# Double elevation: Autonomous weapons and the search for an irreducible law of war

Ioannis Kalpouzos*

City Law School, City, University of London, Northampton Square, London EC1V 0HG; Global Legal Action Network (GLAN), Garden Court Chambers, 57–60 Lincoln's Inn Fields, London, WC2A 3LJ
Email: Ioannis.Kalpouzos.1@city.ac.uk

## Abstract

What should be the role of law in response to the spread of artificial intelligence in war? Fuelled by both public and private investment, military technology is accelerating towards increasingly autonomous weapons, as well as the merging of humans and machines. Contrary to much of the contemporary debate, this is not a paradigm change; it is the intensification of a central feature in the relationship between technology and war: double elevation, above one's enemy and above oneself. Elevation above one's enemy aspires to spatial, moral, and civilizational distance. Elevation above oneself reflects a belief in rational improvement that sees humanity as the cause of inhumanity and de-humanization as our best chance for humanization. The distance of double elevation is served by the mechanization of judgement. To the extent that judgement is seen as reducible to algorithm, law becomes the handmaiden of mechanization. In response, neither a focus on questions of compatibility nor a call for a 'ban on killer robots' help in articulating a meaningful role for law. Instead, I argue that we should turn to a long-standing philosophical critique of artificial intelligence, which highlights not the threat of omniscience, but that of impoverished intelligence. Therefore, if there is to be a meaningful role for law in resisting double elevation, it should be law encompassing subjectivity, emotion and imagination, law irreducible to algorithm, a law of war that appreciates situated judgement in the wielding of violence for the collective.

Keywords: artificial intelligence; autonomous weapons; irreducible law; law of war; targeting

## 1. Introduction

We are experiencing an escalation of both hope and angst in relation to the socially transformative role of technology. Artificial intelligence is gradually colonizing our daily lives as well as our perception of the future. Technological advance informs both the material and economic relations in society, with added investment through the public and private sectors,[1] as well as pervasive intellectual and moral soul-searching. Attitudes range from extreme optimism,[2] or at least utopian

[1]A. Campolo et al., 'AI Now 2017 Report', available at ainowinstitute.org/AI_Now_2017_Report.pdf; V. Boulanin, 'Mapping the Innovation Ecosystem Driving the Advance of Autonomy in Weapon Systems', (2016) *SIPRI Working Paper*.

[2]R. Kurzweil, *The Singularity is Near: When Humans Transcend Biology* (2005). See also the illuminating and entertaining reporting in M. O'Connell, *To Be a Machine: Adventures among Cyborgs, Utopians, Hackers, and the Futurists Solving the Modest Problem of Death* (2017).

engagement,[3] to a range of voices and scholarship highlighting dangers or warning against a catastrophe.[4]

In war, the above spectrum of stances on technology as a whole is both mirrored and intensified. Attitudes range from wild optimism about human perfectibility and the overcoming of human fallibility through artificial intelligence,[5] through pragmatic and productive adaptability,[6] to serious concerns about the de-humanization of war-fighting and the removal of the human beings from the proverbial loop.[7]

The wild swings of technological optimism and pessimism have pedigree. From the biblical figure of the Golem and Hephaistos' robots to Goethe's sorcerer's apprentice, man has fretted over the emancipation of his mechanical creations and the loss of control over his reified wishes, while striving to both physically and intellectually perfect and overcome his humanity through technology.[8] Angst over the brave new world of robots, and its effect on the social, is a recurring affliction.[9]

Now, a prodigious multi-disciplinary literature reckons with artificial intelligence, war and law, and specifically the development and deployment of what are referred to as increasingly, and eventually fully, 'autonomous' weapons. Editorials abound, campaigns are launched to 'stop killer robots',[10] and states assemble to debate their regulation and very existence.[11]

This article intervenes in a rather crowded debate. Why is there such a flurry of current interest in autonomous weapons and their legal regulation? A visceral reaction to the idea of mechanized killing, combined with the perception of the acceleration of artificial intelligence research,[12] may be a sufficient answer. The anticipation of incoming practical problems to-be-solved, questions posed by powerful interested parties, may also be an incentive for lawyers to place their analyses in the marketplace of ideas.

But I believe there is more to it. The present is the time to imagine the future, especially when one is both propelled by and unmoored from the past. The mix of artificial intelligence, war, and law draws some of its headiness from a specific historical moment – a culmination of and a departure from enlightenment rationalism; the apex of progress and the moment when we fear it will get out of hand. 'The end of the end of history'[13] marks a daunting beginning.

The law of war carries this tension, with respect to both the articulation of rules and their enforcement. The legal institution of war rests on the survival of a soldier's individual sense of humanity at the time when his life is laid out for the collective. So does individual (criminal) liability: without that assumption – without that fiction – it makes no sense. I argue that the

---

[3]N. Srnicek and A. Williams, *Demand Full Automation: Postcapitalism and a World Without Work* (2016).

[4]N. Bostrom, *Superintelligence: Paths, Dangers, Strategies* (2014).

[5]R. C. Arkin, *Governing Lethal Behavior in Autonomous Systems* (2009).

[6]K. Anderson and M. Waxman, 'Law and Ethics for Autonomous Weapons Systems: Why a Ban Won't Work and How the Laws of War Can', *Stanford University*, 2013, available at scholarship.law.columbia.edu/faculty_scholarship/1803.

[7]See, for example, Human Rights Watch, 'Losing Humanity: The Case Against Killer Robots', 19 November 2012, available at www.hrw.org/report/2012/11/19/losing-humanity-case-against-killer-robots; P. Asaro, 'On Banning Autonomous Weapons Systems: Human Rights, Automation and the Dehumanization of Lethal Decision-Making', (2012) 94 *International Review of the Red Cross* 687.

[8]O'Connell, *supra* note 2.

[9]Fears of the effects of automation on the job market have featured, for example, on the covers of *Der Spiegel* magazine in March 1964, April 1978, and September 2016.

[10]See www.stopkillerrobots.org.

[11]See, on the 2019 meeting of the Group of Government Experts, established by 2016 Fifth Review Conference of the High Contracting Parties to the Convention on Certain Conventional Weapons (CCW), and related documents, www.unog.ch/80256EE600585943/(httpPages)/5535B644C2AE8F28C1258433002BBF14?OpenDocument.

[12]See the profile and interview of N. Bostrom in R. Khatchadourian, 'The Doomsday Invention: Will artificial intelligence bring us utopia or destruction?', *The New Yorker*, 23 November 2015.

[13]G. Simpson, 'The End of the End of History: Some Epitaphs for Liberalism', (2016) 15(1) *Baltic Journal of International Law* 332–43.

increasing mechanization of warfare pursues the creation of distance from our enemies and from ourselves and reduces the knowledge and intelligence of the law and its application through individual judgement. I further argue that legal research on new weapons technology should focus less on questions of compatibility between given legal rules and the algorithmic and kinetic features of new weapons, as most current scholarship does, and more on an understanding of law as non-reducible to algorithmic engineering.

I start, in Section 2, by reviewing the state of the art-in-the-making. Rather than providing an exhaustive taxonomy I aim to highlight the teleological nature of artificial intelligence research and the industry's investment in the dialectic of increasing machine autonomy and human/machine merging, or 'merged heteronomy'. I then see, in Section 3, this relationship between technology and war in some historical perspective, with a view to discerning functions relevant to law. I argue that the role of technology in war entails a double elevation: above one's enemy and above oneself. The elevation above one's enemy, discussed in Section 3.1, serves a both offensive and defensive impetus and aspires to both spatial and moral/civilizational distance. The elevation above oneself, discussed in Section 3.2, is for self-perfection. It is often associated with a certain understanding of Cartesian dualism and a belief in rational improvement that may see humanity as the cause of *in*humanity and de-humanization as our best chance for humanization. It seeks to mechanize judgement and therefore establish a distance from human failings. It is served by war as governance from a distance and by the increasing physical and cognitive merging of humans and machines. Both physically and, to some extent, in moral and civilizational terms, technology and automation promise such improvement through establishing a distance from the human – our human enemy and our human self. The establishment of this distance entails a decreasing role for human judgement and the weakening of responsibility for such judgement. I further argue, in Section 3.3, that law, or certain strands of mainstream jurisprudence, are complicit in such mechanization, to the extent that law is treated as logic even conceivably reducible to algorithm. And that the reaction against this process of distancing and de-humanization, to the extent that it idealizes the proscriptive or regulatory role of law, is bound to disappoint.

Finally, in Section 4, I aim to begin articulating what the role of law, and legal scholarship, should be in response. I argue that, while responding to a justified angst, the calls for a ban are unlikely to succeed and may miss the target. I turn to long-standing philosophical and sociological critiques which starkly show the limitations of the cognitive theory underpinning artificial intelligence, its disembodied poverty. I argue, and conclude, that if there is to be a meaningful role for law and if we are not to mechanize and outsource our judgement we need to work towards an irreducible and situated understanding of the law of war, one that entails the appreciation of subjectivity and emotion, a law that cannot be coded.

## 2. Towards full autonomy and merged heteronomy

### 2.1. Defining technology in escalation

This is an analysis of the present development of future technology. As such, it can be based on two parameters: the first is the observation of the trajectory of technological development, from the (recent) past to the present, including the present projections of technological expertise; the second relates to 'our beliefs about what it means to be a human being in that future'.[14] In this case as well, and if '[a]rmaments embody fantasies of future conflicts',[15] the discussion of the future is a discussion of the present – our present beliefs, understanding, and projections on humanity and war-fighting. Accordingly, an analysis confined to finding or setting out conditions for

---

[14]G. Noll, 'Weaponising Neurotechnology: International humanitarian law and the loss of language', (2014) 2(2) *London Review of International Law* 201, at 204.

[15]Ibid. See also L. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions* (2007), 1: 'cultural conceptions have material effects'.

'compatibility' of future technology with present law or, conversely, arguing for the 'adaptability' of present law to encompass future technology, would be insufficient for both present and future purposes. To the extent that new weapons technologies reflect an ongoing trend in technology, law, and war and to the extent that they constitute a qualitative leap, discussing them without critically assessing our present categories would be a crucial opportunity lost. Law and technology are in dialogue, in a relationship of mutual influence and this is an already long and established relationship. An argument on the relationship between law and future technology, therefore, while appreciating projected material change, needs to be primarily an argument for the present and how to change, for the future, what is already here.

Weapons technology is described as in escalation towards an ultimate end: full autonomy. Teleological categorization applies, for example, to the qualities of 'adaptiveness'[16] or 'self-governance'[17] and is reflected in the terminology used. At one end, the term 'automatic' describes mechanical response to sensory input, without the ability to adapt to changes in the environment. One step further, the possibility to adapt, but alongside 'a pre-defined set of rules' towards an outcome, is seen to describe 'automated' weapons. Finally, an 'autonomous' weapon, or system,

> is capable of understanding higher-level intent or direction . . . deciding a course of action, from a number of alternatives, without depending on human oversight or control, although these may still be present. Although the overall activity of an autonomous [system] will be predictable, individual actions may not be.[18]

The idea of autonomy dominates the analysis of weapons technology and its evolution.

The increase in the autonomy of weapons is usually perceived as corresponding to the concomitant decrease of the role of individuals. The terminology of the human/machine command and control relationship accordingly ranges from the semi-autonomous or human *in* the loop (where human input is *required*), through supervised autonomy or human *on* the loop (where an individual can intervene when something goes wrong), to fully autonomous or human *out of* the loop. Here the weapons systems 'operate completely on their own and . . . humans are not in a position to intervene'.[19]

The evolution of autonomy is piecemeal, a process of filling in gaps, an increasing accumulation of different skills towards full functional and operational autonomy. In the meantime, it is possible for a weapon to have full autonomy in terms of identifying and engaging a target but no autonomy in kinetic terms. Within a specific task, even if a weapon has full autonomy in, for example, identification and engagement, this may be supported by a very basic level of cognitive sophistication. A landmine, indeed, satisfies the above examples, providing no kinetic autonomy, full engagement autonomy, and very limited cognitive capacity.[20]

When it comes to the representation and application of behavioural, and legal, rules, cognitive autonomy and the capacity to set goals is seen as crucial. For Sartor and Omicini 'only teleological systems can be fully endowed with the capacity to be guided by norms, as elements that play a specific role in the deliberative process of such systems'.[21] Such independent, adaptive, and

---

[16]See G. Sartor and A. Omicini, 'The autonomy of technological systems and responsibilities for their use', in N. Bhuta et al., *Autonomous Weapons Systems: Law, Ethics, Policy* (2016), 39, at 49.

[17]'[T]he degree of autonomy is often measured by relating the degree at which the environment can be varied to the mean time between failures, and other factors indicative of robot performance.' S. Thrun, 'Toward a framework for human-robot interaction', (2004) 19(1) *Human-Computer Interaction* 9–24, at 14.

[18]D. Mindell, *Our Robots, Ourselves: Robotics and the Myths of Autonomy* (2015), 12.

[19]Boulanin, *supra* note 1, at 12. See also N. Sharkey, 'Staying in the loop: human supervisory control of weapons', in Bhuta et al., *supra* note 16, at 27 for a useful five-part categorization.

[20]Sartor and Omicini, *supra* note 16, at 44–8.

[21]Ibid., at 52.

purposeful machine agents can exist independently or within 'artificial agent societies',[22] some-times referred to as 'swarms'.[23]

'Fully autonomous systems' then, in relation to war-fighting, are systems that, once deployed, are able to adapt, receive, and process feedback, and display a level of functional autonomy that effectively does not distinguish them from human decision makers.[24] If anything, in fact – and that is really the point – such systems may be capable of a higher level of tactical or even strategic decision-making (cognitive) and war-fighting (kinetic) capacity. This seems to be a shared under-standing among states and NGOs, otherwise holding different positions in the autonomous weap-ons debate. Accordingly, the US Department of Defense refers to '[a] weapons system that, once activated, can select and engage targets without further intervention by a human operator'.[25] Human Rights Watch, while setting out its position against autonomous weapons systems, defines them as '[r]obots that are capable of selecting targets and delivering force without any human input or interaction'.[26] Other states[27] and organizations[28] provide similar definitions. These func-tions of autonomy, and their escalation, entail both physical and cognitive distancing from human agents in the overall process of targeting.

Although the concept and image of autonomy dominates the discourse in a way that influences, as we will see in the next section, much of the scientific research and development, it only partially describes technological escalation. In fact, the artificial intelligence of war-fighting complements increasing autonomy with what has been called 'merged heteronomy'. Increasing cognitive auton-omy, and distance, coexists with increasingly close physical proximity. This serves pragmatic aims: Technological limitations in the fragmented development of different aspects of autonomy, the continuing necessity of human input, and the fragility of human/machine networks mean that the individual's continuing presence in the loop remains an operational necessity. This also means that a full-on confrontation with social and political resistance to the reality of distinct *killer robots* is placed in abeyance. Continuous, if vague, assurances of the human remaining *in the loop* and retaining *meaningful control* are thereby facilitated.

And yet, the abeyance is a trap and our presence 'in the loop' is no guarantee. To the extent that 'full autonomy' is understood to require the separate physical existence of an, often anthropomor-phized, robot other, it obscures the crucial role that increasing cognitive autonomy plays in a nom-inally heteronomous decision-making process. Cognitively, as well as physically and kinetically, humans and machines become decreasingly separate, less and less other. Their understanding of the rules, the *nomos*, is merging, as is their physical existence. Increasing autonomy and merged heteronomy serve the same purpose, the same teleology of mechanization. 'The loop', itself, is changing, increasingly relying on artificial intelligence.

[22]See, e.g., S. Ossowski, *Co-ordination in Artificial Agent Societies: Social Structures and Its Implications for Autonomous Problem-Solving Agents* (1999).

[23]See P. Scharre, *Robotics on the Battlefield Part II: The Coming Swarm* (2014); M. Rubenstein et al., 'Programmable Self-Assembly in a Thousand-Robot Swarm', (2014) 345 *Science* 795.

[24]See also C. Heyns, 'Autonomous weapons systems: living a dignified life and dying a dignified death', in Bhuta et al., *supra* note 16, 3, at 4; J. D. Ohlin, 'The Combatant Stance: Autonomous Weapons on the Battlefield', (2016) 92 *International Law Studies* 1 on 'functional autonomy'. The idea of functional autonomy of course hails from Turing's 'imitation game'. See A. M. Turing, 'Computing Machinery and Intelligence', (1950) *Mind: A Quarterly Review of Psychology and Philosophy* 433.

[25]US Department of Defence Directive 3000.09, 21 November 2012, available at www.hsdl.org/?abstract&did=726163, Glossary.

[26]Human Rights Watch, 'Losing Humanity: The Case Against Killer Robots', *supra* note 7.

[27]Examination of various dimensions of emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and the purposes of the Convention, submitted by the Netherlands, CCW/GGE.1/2017/WP.2, 9 October 2017.

[28]Geneva Academy, 'Autonomous Weapon Systems under International Law' (Academy Briefing no. 8, November 2014), at 6, available at www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf.

## 2.2. The state of the art-in-the-making

The escalating aspirations of full autonomy, in combination with merged heteronomy, can be seen in both existing and projected weapons technology.[29] At this stage, increasing autonomy is more confidently deployed and developed in defensive weapons systems or in surveillance and evidence gathering technology. There are operational weapons technologies of a defensive nature that employ a level of autonomous decision-making.[30] An example of an existing weapon displaying a significant degree of automation is the Super aEgis II, an anti-personnel sentry weapon system manufactured by South Korean DoDAMM. The turret gun uses thermal imaging to offer an autonomous detection, tracking, and targeting capacity of vehicle or human targets within a three kilometre range.[31] The weapon, currently operating in the Korean Demilitarized Zone, has the option of operating on a fully automated mode, although it is currently held on a 'slave' mode, with an individual in the loop.[32]

Offensive weapons employing elements of autonomy range from long-range anti-ship smart missiles (LRASM) to the Harpy loitering munitions. The LRASM, manufactured by Lockheed Martin,[33] is a long-range, precision-guided anti-ship missile, which may chart its course both in accordance with pre-routing as well as autonomously, in order 'to find and destroy its predetermined target in denied environments'.[34] Loitering munitions are disposable unmanned aerial vehicles (UAVs), known as 'kamikaze drones', targeted at an overall area, where they loiter until they can find and strike specific ground targets.[35] While most are currently operated by human agents who 'close the (sensor-to-shooter) circuit and hit the target',[36] according to the Head of Israeli Aerospace Industries' Land Systems Division, current technology 'may be operated without human involvement, and such involvement will only depend on the fire employment guidelines that are based on non-technological considerations'.[37]

Ongoing research and development aimed at increasing autonomous kinetic ability is crucially complemented by the investment in the capacity to survey, identify, and engage potential targets. Project Maven[38] introduced artificial intelligence and machine learning innovations to intelligence, surveillance, and target acquisition and integrated it into the battlefield. Massive amounts of data, the product of drone surveillance, are analysed for the identification of objects and potential targets. Machine learning sprints are developing the algorithm.[39] This allows both the classification of images for the US military and the rapid improvement of the program. While Project Maven spokespeople assuage concerns by confirming that individuals are the ones reviewing the

---

[29] For a number of examples see D. Lewis et al., *War-Algorithm Accountability* (2016), at 34.

[30] See generally Boulanin, *supra* note 1, at 26.

[31] See www.dodaam.com/eng/sub2/menu2_1_4.php.

[32] See also Samsung's SGR-1 sentry robots at www.defensereview.com/samsung-sgr-a1-armedweaponized-robot-sentry-or-sentry-robot-remote-weapons-station-rws-finally-ready-for-prime-time/.

[33] See G. Reim, 'Lockheed Martin delivers first Long Range Anti-Ship Missiles', *Flight Global*, 20 December 2018, available at www.flightglobal.com/news/articles/lockheed-martin-delivers-first-long-range-anti-ship-454597/.

[34] See www.lockheedmartin.com/us/products/LRASM/overview.html and www.youtube.com/watch?v=h449oIjg2kY.

[35] Boulanin, *supra* note 1, at 50–55. See also D. Brennan, 'Suicide Drones: Are Tiny Missiles That "Loiter" in the Air for Hours the Future of Assassination Wars?', *Newsweek*, 22 February 2019, available at www.newsweek.com/drones-suicide-kamikaze-war-assassination-missile-uav-war-1340751.

[36] See A. Rapaport, 'Loitering Munitions alter the Battlefield', *IsraelDefense*, 30 June 2016, available at www.israeldefense.co.il/en/content/loitering-munitions-alter-battlefield.

[37] See www.israeldefense.co.il/en/node/31130.

[38] C. Pellerin, 'Project Maven Industry Day Pursues Artificial Intelligence for DoD Challenges', *Department of Defence News*, 27 October 2017, available at www.defense.gov/News/Article/Article/1356172/project-maven-industry-day-pursues-artificial-intelligence-for-dod-challenges/.

[39] See Deputy Secretary of Defense's Memorandum on the Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven), 26 April 2017, available at www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf; D. Lewis, N. Modirzadeh and G. Blum, 'The Pentagon's New Algorithmic-Warfare Team', *Lawfare*, 26 June 2017, available at www.lawfareblog.com/pentagons-new-algorithmic-warfare-team.

algorithms' classifications and selecting the potential targets and that Maven has not been used for specific targeting decisions, the algorithms are tested live, integrating the combat theatre, rather than in a lab environment. In the combat theatre, presumably, computer identified objects are actioned. At the same time the algorithm is constantly learning, increasingly ready for fuller autonomy.[40]

Maven was just the beginning. The newly released US Department of Defense Artificial Intelligence Strategy[41] is creating a new Joint Artificial Intelligence Center, a Department priority,[42] headed by Lt. Gen. Jack Shanahan, head of Project Maven. The aims of Maven are at the heart of current US research, which is focusing especially on image analysis,[43] 'improving the capabilities of sensing algorithms for autonomous surveillance and targeting',[44] including through stealth technology for UAVs, enabling them to operate autonomously in 'communication-denied airspace' for the purposes of both surveillance and targeting. The image recognition automation is not limited to non-human objects, but includes focus on facial recognition software, some research focusing on 'probabilistic algorithms that determine the likelihood of adversarial intent',[45] reflecting the increasing influence of a criminal law paradigm on the law of targeting.

Similar trends may be observed in human/machine technologies. Kinetic autonomy, such as 'Fast Lightweight Autonomy',[46] is combined with the development of natural language processing for human-machine communication.[47] Cognitive autonomy, such as the 'probabilistic programming for advanced machine learning',[48] is sought alongside the ability of artificial intelligence systems to 'explain themselves' and earn the trust of human beings;[49] and the collaboration of autonomous agents amongst themselves,[50] or under the control of a reduced number of human operators.[51] 'Swarm squadrons of network enabled drones', are, according to the former UK Defence Secretary, part of 'the future direction of the UK armed forces'.[52] Such research aspires to achieve the crucial goal of strengthening network contact in complex human/machine systems, towards their further integration.

---

[40]The Department of Defense's Algorithmic Warfare Cross-functional Team (AWCFT)'s 'objective is to turn the enormous data available to DoD into actionable intelligence and insights at speed'. See Deputy Secretary of Defense's Memorandum on the Establishment of an Algorithmic Warfare Cross-Functional Team, ibid.

[41]Summary of the 2018 US Department of Defense Artificial Intelligence Strategy, released 12 February 2019, available at media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

[42]Memorandum on the Establishment of Joint Artificial Intelligence Centre, 27 June 2018, available at admin.govexec.com/media/establishment_of_the_joint_artificial_intelligence_center_osd008412-18_r . . . .pdf.

[43]US Department of Defense AI Strategy, at 11. See also, on the priority of the automation in geolocation, M. Ekelhof, 'Lifting the Fog of Targeting: "Autonomous Weapons" and Human Control through the Lens of Military Targeting', (2018) 71(3) *Naval War College Review*, 61, 80 and footnote 107 (reporting an interview with NGA technical director and big-data specialist).

[44]Boulanin, *supra* note 1, at 17. DARPA's TRACE program is described thus: 'The Target Recognition and Adaption in Contested Environments (TRACE) program seeks to develop an accurate, real-time, low-power target recognition system that can be co-located with the radar to provide responsive long-range targeting for tactical airborne surveillance and strike applications.' See www.darpa.mil/program/trace. See also the modestly called 'Imaging Through Almost Anything Anywhere (ITA3)' program at www.darpa.mil/program/fast-lightweight-autonomy.

[45]See the press release at www.modusoperandi.com/modus-operandi-awarded-1-million-u-s-army-contract-for-enemy-and-criminal-behavioral-recognition-system/.

[46]'[T]he program aims to develop and demonstrate the capability for small (i.e., able to fit through windows) autonomous UAVs to fly at speeds up to 20 m/s (45 mph) with no communication links to the operator and without GPS guidance.' at www.darpa.mil/program/fast-lightweight-autonomy.

[47]See DARPA's project 'Communicating with Computers' (CwC), at www.darpa.mil/program/communicating-with-computers.

[48]www.darpa.mil/program/probabilistic-programming-for-advancing-machine-Learning.

[49]See www.darpa.mil/program/explainable-artificial-intelligence.

[50]See the Micro Autonomous Systems Technology (MAST) research program at www.mast-cta.org/.

[51]Boulanin, *supra* note 1, at 17.

[52]For the UK Defence Secretary's statement on developing swarms see www.gov.uk/government/speeches/defence-in-global-britain.

Finally, while the Tactical Assault Light Operator Suit (TALOS) project, colloquially referred to as the Iron Man suit, failed, individual components will be used[53] and it represents a clear, if spectacular, statement as to the intended future of heteronomy merging so completely that the distinction between human and machine will increasingly disappear. The suit would be a computerized exoskeleton that would increase both physical and cognitive performance offering 'increased survivability, lethality, situational awareness and decreased time to target engagement'.[54] The development of Brain-Computer Interface is at the heart of DARPA's latest call for an Intelligent Neural Interfaces program, aiming at 'modeling and maximizing the information content of biological neural circuits to increase the bandwidth and computational abilities of the neural interface'.[55] Cognitive enhancement will be achieved by integrating human and artificial intelligence.[56]

The dialectic of autonomy and merged heteronomy is supported by powerful socioeconomic forces. The new US Department of Defense artificial intelligence strategy expressly, and insistently, seeks to integrate both academic and commercial actors in the development of future weapons technology.[57] Embracing artificial intelligence is seen as a holistic national, social and economic endeavour; a cultural aspiration.[58] The relationship of Google with Project Maven is indicative of the enthusiasm, the tension, and the eventual 'synergy' between the military and private commercial actors. An initial embrace, and the license for corporations to own the intellectual property of the improved algorithm, led to a high-profile employee reaction and Google's divestment,[59] while the future relationship remains open.[60] While the US has been the most transparent, or even outspoken,[61] the public/private model that the US has pioneered is being emulated in, for example, Russia, China,[62] and Turkey.[63]

States' positions on the degree of weapons emancipation reflect the tension between the technological urge and remaining taboos. They are, accordingly, somewhat vague or open to change. The current US position is set out in the Department of Defense Directive 3000.09 which requires autonomous weapons systems to have the 'capability to allow commanders and operators to exercise appropriate levels of human judgment in the use of force ...'[64] and sees the research described above, including Project Maven, as below this threshold.[65] The UK has stated that

---

[53]See www.defenseone.com/technology/2019/02/us-military-chopping-its-iron-man-suit-parts/154706/.

[54]See www.socom.mil/sof-at-l-seeks-industry-partners-for-future-talos-innovations.

[55]See www.fbo.gov/index?s=opportunity&mode=form&id=f48385e64664e78b061032cd12b86db5&tab=core&_cview=0.

[56]See also Y. Sankai and T. Sakurai, 'Exoskeletal cyborg-type robot', (2018) 3(17) *Science Robotics*, available at robotics.sciencemag.org/content/3/17/eaat3912.

[57]US Department of Defense AI Strategy, at 12.

[58]The Deparment of Defense's reference to its own culture reflects a broader culture of business and risk. See US DoD AI Summary, at 14: 'We are building a culture that welcomes and rewards appropriate risk-taking to push the art of the possible: rapid learning by failing quickly, early, and on a small scale.'

[59]D. Wakabayashi and S. Shane, 'Google Will Not Renew Pentagon Contract That Upset Employees', *New York Times*, 1 June 2028, available at www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

[60]L. Fang, 'Google hedges on promise to end controversial involvement in military drone contract', *The Intercept*, 1 March 2019, available at theintercept.com/2019/03/01/google-project-maven-contract/.

[61]See Boulanin, *supra* note 1 and Appendix C for a list of corporations.

[62]See E. B. Kania, 'Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power', *Center for a New American Security*, 28 November 2017, available at www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

[63]See www.dailysabah.com/war-on-terror/2017/10/02/military-to-utilize-artificial-intelligence-for-counterterrorism.

[64]US Department of Defense Directive, *supra* note 32. For the 'troubling lacunae' of the Directive, see D. Saxon, 'A human touch: autonomous weapons, DoD Directive 3000.09 and the interpretation of "appropriate levels of human judgment over the use of force"', in Bhuta et al., *supra* note 16, at 185.

[65]This is reiterated, but not significantly elaborated, in the Department of Defense's AI strategy, at 15. See also the interview with Shanahan in M. Ekelhof, 'Lifting the Fog of Targeting: "Autonomous Weapons" and Human Control through the Lens of Military Targeting', (2018) 71(3) *Naval War College Review*, 61, 85 and fn. 131 (where he allows that 'it could very well be that, if a major conflict arises, all bets will be off, with states feeling forced into more reliance on autonomous systems because their adversaries are willing to take more risk').

its current research will make sure that individuals remain 'in the loop'.[66] Other states are unapologetic in allowing themselves flexibility.[67]

What we are witnessing is the gradual identification and assembling of different aspects of autonomous capacity, while, with 'full autonomy' in abeyance, human judgement 'in the loop' is increasingly mechanized through human-machine merging. Both parts of this dialectic contribute to the mechanization and distancing of the decision-making process that involves legal judgement. This distancing, which I call 'double elevation', will now be placed in its historical perspective, with a view to begin thinking the future role of international law.

## 3. Double elevation and the distancing of judgement

Technology is not neutral.[68] Assuming the neutrality of technology – and attaching to it the assumed neutrality of law – precludes any critical understanding of either. Technology, its development and use, reflects both theoretical and practical commitments: it 'is covert philosophy'.[69] It exists in a relationship of co-production with culture, politics, and law.[70] As Paul Edwards put it in a seminal study of Cold War weapons technology, 'we can make sense of the history of computers as tools only when we simultaneously grasp their roles as metaphors in . . . the period's . . . science, politics and culture'.[71]

In this section, I argue that the promise of technological progress and automation in war, as in general, is a promise of civilization, a promise of improvement. It entails a double elevation: above one's enemy and above one's self. At the centre of it there is a paradoxical assumption, namely that the non-human can be more humane than the human. The elevation above one's enemy combines military distance with a perception of civilizational and moral superiority. The elevation above oneself aims at creating a distance from human features perceived as weak or unreliable. Both full autonomy and merged heteronomy require the increasing mechanization of human judgement. What we have learned to understand as the civilization of war-fighting rests on and pursues its mechanization.

### 3.1. Rising above one's enemy

Technological distancing aims at developing asymmetry and invulnerability and elevating oneself above one's enemy in both strictly speaking military and broader civilizational terms. The latter

---

[66]The UK ministry of defence stated: 'UK policy is that the operation of weapons will always be under control as an absolute guarantee of human oversight, authority and accountability. The UK does not possess fully autonomous weapon systems and has no intention of developing them.' See M. Savage, 'Humans will always control killer drones, says ministry of defence', *The Observer*, 10 September 2017, available at www.theguardian.com/politics/2017/sep/09/drone-robot-military-human-control-uk-ministry-defence-policy. However, a previous commander of the UK Joint Forces Command has expressed his scepticism that such pledges will be maintained. See B. Farmer, 'Prepare for rise of "killer robots" says former defence chief', *Daily Telegraph*, 27 August 2017, available at www.telegraph.co.uk/news/2017/08/27/prepare-rise-killer-robots-says-former-defence-chief/.

[67]See the Russian statement at www.unog.ch/80256EDD006B8954/(httpAssets)/B7C992A51A9FC8BFC12583BB00637BB9/$file/CCW.GGE.1.2019.WP.1_R+E.pdf, at 4–5.

[68]P. Vermaas et al., *A Philosophy of Technology: From Technical Artefacts to Sociotechnical Systems* (2011), at 16 refer to the well-known National Rifle Association slogan: 'Guns don't kill people. People kill people' as an example of a 'succinct way of summarising what is known as the neutrality thesis of technical artefacts'. See M. Heidegger, 'The Question Concerning Technology', in *Basic Writings* (2008 [1954]), 217, at 217: 'Everywhere we remain unfree and chained to technology, whether we passionately affirm or deny it. But we are delivered over to it in the worst possible way when we regard it as something neutral.'

[69]P. Agre, *Computation and Human Experience* (1997), at 240.

[70]See, more broadly, S. Jasanoff (ed.), *States of Knowledge: The Co-Production of Science and the Social Order* (2004); S. Jasanoff, 'Technology as a Site and Object of Politics', in R. E. Goodin and C. Tilly (eds.), *The Oxford Handbook of Contextual Political Analysis* (2006), 745.

[71]P. Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America* (1997), at ix.

type of elevation allows not simply a geographical distance but also a moral distance with significant consequences for the role of law and judgement in killing.

Military technology is central to early imperialist expansion[72] and its concomitant civilizational pretension, culminating in the steep military and moral asymmetry achieved in nineteenth century colonial warfare. Churchill's description in the context of the Boer war of the British infantry 'steadily and solidly' firing against the Sudanese Dervishes in 'the most signal triumph ever gained by the arms of science over barbarians', while 'the mere physical act became tedious',[73] is illustrative. Technology allows military superiority, guaranteeing the physical safety and invulnerability of one's forces; the asymmetry achieved reflects an already assumed civilizational distance which allows a moral dissociation from the act of killing, expressed in the ennui of physical exertion; the civilization of the technologically advanced party is enforced.[74]

The role of military technology in the elevation above one's enemy is most closely associated with the growth of air power and the aspirations of invulnerability associated with it. Air power, especially in situations of colonial asymmetry, constituted a relationship of vertical distance, allowing the surveillance and policing of one's inferior enemy, both at initial conquest and through the protracted practice of colonial administration and pacification.[75] That colonial relationship achieved new technological heights in the context of the Cold War. Towards the end of the 1960s, the Vietnam impasse pushed for the assertion of asymmetry through the development of an automated battlefield to improve targeting capacity and protect American soldiers. Operation Igloo White attempted the surveillance of the Ho Chi Minh Trail in Laos through the use of camouflaged sensors designed to detect different types of human activity, including body heat, vehicle noise or the smell of human urine,[76] or sweat.[77] When picked up, such activities appeared on the screens in the headquarters' terminals in Thailand and fed into the targeting system of military aircraft. A 'kill box'[78] was constructed and targeted. The operation's centralized, computerized, automated method of 'interdiction' relied on an active global defence and aspirations for the full automation of the battlefield. These are set out by General William Westmoreland, the Chief of Staff of the US Army at the time, in a tenor strongly evocative of our present debate:

> On the battlefield of the future, enemy forces will be located, tracked, and targeted almost instantaneously through the use of data links, computer assisted intelligence evaluation, and automated fire control. . . . I see battlefields on which we can destroy anything we locate through instant communications and the almost instantaneous application of highly lethal firepower. . . . [A]n improved communicative system . . . would permit commanders to be continually aware of the entire battlefield panorama down to squad and platoon level . . . I am confident the American people expect this country to take full advantage of its technology - to welcome and applaud the developments that will replace wherever possible the man with the

---

[72]G. Parker, *The Military Revolution: Military Innovation and the Rise of the West 1500-1800* (1996).

[73]W. Churchill, 'The River War: An Account of the Reconquest of the Soudan (1899)', in D. R. Headrick, *The Tools of Empire* (1981), 118.

[74]See also T. Asad, *On Suicide Bombing* (2007), 34: 'The modern Western army is concerned with engaging efficiently with dangerous, because underdeveloped, peoples, in ways that are at once ruthless and humane, in which brutal attack may become a civilizing sign.'

[75]C. Munro, 'Mapping the Vertical Battlespace: Towards a legal cartography of aerial sovereignty', (2014) 2(2) *London Review of International Law* 233–61; D. Gregory, 'From a View to a Kill: Drones and Late Modern War', (2011) 28 *Theory, Culture & Society* 7–8, 188; S. Moyn, 'Drones and Imagination: A Response to Paul Kahn', (2013) 24(1) EJIL 227.

[76]A. Jaubert, 'Zapping the Viet Cong by computer', (1972) *New Scientist*, 30 March 1972, at 685, 687, available at books.google.co.uk/books?id=juOOP4nRFrQC&lpg=PP1&hl=EN&pg=PP1#v=onepage&q&f=false.

[77]Ibid., at 688.

[78]See G. Chamayou, *Drone Theory* (2015), Ch. 6.

machine . . . With cooperative effort, no more than 10 years should separate us from the automated battlefield.[79]

As it turns out, Operation Igloo White was a complete failure.[80] And yet the technological ambition remained. In 1973 the *New Scientist* echoed General Westmoreland's technological/military optimism. There was 'at present, great interest in the development of remotely piloted vehicles (RPV's) for missions such as reconnaissance, electronic warfare, ground attack and air-to-air combat'.[81] Increasingly, to these purposes was added another: targeted assassination.

The ambition of the precision of a self-sustaining intelligence/targeting loop in drone warfare illustrates the confluence of offensive and defensive imperatives in elevating oneself above one's enemy. Markus Gunneflo has shown how the practice of and legal justification for targeted killings was developed in Israeli and US policy as a means of constitutional protection of the citizens to be distinguished from unlawful assassination.[82] In such 'active defence', especially when exercised globally, we see the merging of the offensive distance of air power, seeking to impose a vertical relationship of war, and the defensive distance of integrated human/machine surveillance systems.

This vision is reflected in the prioritization in the 1990s of drone research.[83] Drones, both the surveillance and the targeting kind, have been seen as symbolizing a 'change of paradigm' in the conduct of war.[84] They, however, follow the trajectory discussed – that of achieving an elevation above one's enemy, associated with geographical distancing and the moral/civilizational distance associated with governing through war from above.[85] The present ambition, of both escalated weapon emancipation and human/machine merging, follows that same path. However autonomous, further distancing remains the goal.[86] This is not a paradigm change.[87] However, to the extent that there is a rapid acceleration of technological development we could, perhaps, refer to an 'avalanche': 'when conditions are ripe, individual events, even small ones, can trigger a massive, downward rush'.[88] This metaphor may serve to describe a well-established trajectory combined with the feeling that things may be spiralling out of control.

From colonial asymmetry to the post-Cold War fighting of 'terror', the elevation above one's enemy through weapons technology guarantees physical and moral distance; it also denotes, and imposes, the pretention of a higher civilization. As we will see, the promise of precision, professionalization, optimization of decision making – with humans involved, but assisted by technology – underlines another kind of elevation: one that supposedly saves humans from themselves.

---

[79]Address by General W. C. Westmoreland, Chief of Staff, US Army, Annual Luncheon Association of the United States Army, Sheraton Park Hotel, Washington, DC, 14 October 1969 (*Congressional Record,* US Senate, 16 October 1969).

[80]See Edwards, *supra* note 71, at 7.

[81]F. Barnaby, 'Towards tactical infallibility', *New Scientist*, 10 May 1973, 348–54, at 351.

[82]M. Gunneflo, *Targeted Killing: A Legal and Political History* (2016).

[83]A. Cockburn, *Kill Chain: Drones and the Rise of High-Tech Assassins* (2016).

[84]For their 'mythical' role in the production of a 'new paradigm' of law and war see I. Kalpouzos, 'The Armed Drone', in J. Hohmann and D. Joyce (eds.), *International Law's Objects* (2018).

[85]Munro, *supra* note 75; Gregory, *supra* note 75.

[86]DARPA's TRACE program sets out the logic quite clearly, using language which connects Vietman's electronic battlefield, through drone use, towards escalated automation, while highlighting the problematic effects of distance on the network's reliability thus requiring the strengthening and further integration of a human/machine system: 'In a target-dense environment, the adversary has the advantage of using sophisticated decoys and background traffic to degrade the effectiveness of existing automatic target recognition (ATR) solutions. Airborne strike operations against relocatable targets require that pilots fly close enough to obtain confirmatory visual identification before weapon release, putting the manned platform at extreme risk. Radar provides a means for imaging ground targets at safer and far greater standoff distances; but the false-alarm rate of both human and machine-based radar image recognition is unacceptably high', available at www.darpa.mil/program/trace.

[87]The, arguably overused, phrase, hails from T. Kuhn's, *The Structure of Scientific Revolutions* (1962) where he argues that instead of viewing science as a rational cumulative process, it should be understood as entailing 'intellectual revolutions' where 'one conceptual world is replaced by another' (p. 10).

[88]L. Daston and P. Galison, *Objectivity* (2007), 49.

### 3.2. Rising above oneself

'We are in an arms race with ourselves – and we are winning'[89]

Technological evolution in war is not only about overcoming the enemy. It is also about overcoming one's own imperfections in the wielding of violence. It is a process of progress, improvement, rationalization, optimization, ultimately the civilization of war-fighting. The role of this second elevation, which both facilitates and aims to justify the elevation above one's enemy, is often underappreciated. I will highlight it in this section, complementing the historical narrative above and recognizing its influence on a certain view of the relationship between technology, war and international law.

Elevation above oneself does not require asymmetrical relationships. The technological impetus of air power did not only serve the purpose of offense. It played a crucial role in the development of the relationship between human and machine for defensive purposes. The efforts to counter distancing and provide an effective defence against the German Luftwaffe and the early smart bomb technology of the V-1 and V-2 missiles significantly pushed forward artificial intelligence research.[90] One such effort, led by Norbert Wiener, focused on the scientific articulation of human-machine interaction and the understanding of a pilot and his aircraft as a single unit, an integrated system, the behaviour of which could be predicted. While not successfully weaponized, the research led to Wiener's theory of cybernetics,[91] a widely influential theory for the scientific understanding of information, communication and the function of individuals in their socio-technical environment.

Cybernetics is crucial for the evolution of human/machine merging, and the perception of self-improvement alongside the elevation above one's adversary. Cybernetics is especially important for re-thinking law and agency in autonomous systems as it is, at the same time, based on a formalized understanding of information as the elementary unit for any sort of communication (human/human, human/machine, machine/human, or machine/machine) while having critical implications in relation to our understanding of agency and autonomy in human/machine systems. Therefore, it can be useful in appreciating that increasing autonomy and merged heteronomy are not opposites and that a 'human-in-the-loop' is not, by itself, the answer to the question of mechanization of judgement.[92]

For cybernetics, as Peter Galison has pointed out, the enemy, the German Luftwaffe, with its smart missiles and able pilots, is already perceived as hyper-rational, an advanced unit of human/machines, 'a mechanized Enemy Other'.[93] The distance already achieved by the enemy is an impetus for understanding them, through Wiener's research, as a merged human/machine system. This perception of the enemy and the effort to predict their behaviour extends to and corresponds to the cybernetic perception of the world, and ourselves, as merged human/machine systems. The understanding of the enemy's humanity as partial, as merged with a technical system, is reflected back to the view of oneself and it is emulated. Their distance becomes our distance; their elevation

---

[89]J. Wiesner, chairman of the Science Advisor Committee to President John F Kennedy, quoted in G. Allison and F. Morris, 'Armaments and Arms Control. Exploring the Determinants of Military Weapons', in F. Long and G. Rathjens (eds.), *Arms, Defence Policy, and Arms Control* (1976), at 119.

[90]T. Rid, *Rise of the Machines: The Lost History of Cybernetics* (2017), Ch. 1.

[91]N. Wiener, *Cybernetics: or control and communication in the animal and the machine* (2013 [1948]).

[92]See also M. Arvidsson, 'Targeting, Gender, and International *Posthumanitarian* Law and Practice: Framing the Question of the Human in International Humanitarian Law', (2018) 44(1) *Australian Feminist Law Journal* 9. See also para. 22 of Annex III of the 2018 GGE Report where the 'human' is the only stable, and unproblematized, parameter in the discussion of meaningful human control over autonomous weapons.

[93]P. Galison, 'The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision', (1994) 21(1) *Critical Inquiry* 228. Galison points out, at 231, that, alongside cybernetics, this idea of a calculating enemy also motivated the development of game theory.

is the impetus for ours. A formalized system of information sharing and a merging in technological structures is the way forward. This is what cybernetics endeavoured to provide.

Such impetus for self-improvement through military technology was applied to the creation of broader systems for the governance of war. In the Cold War, alongside the offensive asymmetry of Vietnam's aspired automated battlefield, the period saw the creation of sophisticated human/machine systems for defensive purposes as well. The massive investment in the Semi-Automatic Ground Environment (SAGE) system in the 1950s – 'the first large-scale, computerized command, control and communications system' was aimed at 'global oversight and instantaneous military response'.[94] The identification of and response to the incoming threats would remain at a distance, achieved through the merging of human and machine surveillance power, in a complex and holistic system of artificial intelligence.

In this mode of active defense, elevation above oneself and elevation above one's enemy are seen as mutually reinforcing. The creation of distance and asymmetry in the elevation above one's enemy envisions the conduct of war through an increasingly vertical relationship akin to governance.[95] This entails qualities and aspirations associated with rational governance.[96] Such qualities, like the rationalization and optimization of decision-making, span the range of war-making, from the level of planning and prioritizing targets (for example, on the production of 'kill lists'[97]) to the level of the individual decision maker: the one pressing the button. The self-improvement through technology that puts one party in a position to govern through war is displayed in *how* that party governs through war, justifying its dominance.

Elevation above oneself through technology is not, of course, limited to the conduct of war. Progress and self-improvement through technology is inscribed in a particular narrative of civilization, evolution and progress. Historical investment in technology has at the same time aimed at the realization of human potential and the transcendence of human limitations.[98] In this sense, it is a metaphysic. It both celebrates humanity and aims to move beyond it.

This tradition of thought will be engaged with, and related to law, in some more detail in the next section. Here, we will recognize its influence on the way technology is seen as serving the humanization of war and law. The belief in the improvement of and on humanity can be observed on two levels: Firstly, technology is believed to be a progressive force due to its effects: in the bettering of the conditions of life (or the modalities of killing). Secondly, belief in the salutary and transcending effects of technology is associated and credited to a particular way of thinking,[99] believed to have enabled technological progress in the first place. To the extent that technological progress produces thinking machines, this way of thinking is reified, hardwired, embodied in the technology itself – and fed back to human beings who interact with the machines they have created. The first level of belief in technology can be seen in the context of the discussion of the weapons' effects. The second one is especially relevant to how machines and humans interact with law.

Starting with the former, both distancing and the promise of precision[100] associated with technology – from smart bombs, to drone surveillance/targeting, to algorithmic target selection – are often perceived as allowing for higher levels of discrimination in targeting. It may be that results

---

[94]Edwards, *supra* note 71, at 75.

[95]M. Valverde and M. Lomas, 'Insecurity and Targeted Governance', in W. Larner and W. Walters (eds.), *Global Governmentality* (2004), 233, at 245.

[96]M. Weber, 'The Vocation of Science', in S. Whimster (ed.), *The Essential Weber: A Reader* (2004), 270.

[97]See J. Weber, 'Keep adding. On kill lists, drone warfare and the politics of databases', (2016) 34(1) *Society and Space* 107–25.

[98]See O'Connell, *supra* note 2, at 142: 'If we want to be more than mere animals, we need to embrace technology's potential to make us machines.'

[99]See P. Husbands et al., 'Introduction: The Mechanical Mind', in P. Husbands et al. (eds.), *The Mechanical Mind in History* (2008).

[100]On the function of this promise in relation to the object of the drone, see Kalpouzos, *supra* note 84.

on the ground challenge such promises,[101] perhaps partly due to the license that users of advanced military technology felt able to take in setting out their input parameters.[102] However, the promise remains, and the development of targeting technology is seen to contribute to the humanization of war. New weapons, and imagined future weapons all the more, are seen as promising a level of precision heretofore unprecedented. This feeling is shared, and expressed, both by governments[103] and scholars.[104] While current technological limitations, for example in terms of the limitations of face recognition technology, are often conceded, the general trajectory of humanization through precision is repeatedly asserted. Indeed, such precision is identified at various levels, including the identification of kill lists,[105] the taking of precautions,[106] and the launching of attacks,[107] to the extent that scholars even talk of a future *obligation* to use autonomous weapons systems.[108] Technology allows certainty and predictability.[109]

Moreover, as a matter of practice and individual decisions, increasing autonomy and merged heteronomy are seen to contribute to the elimination of mistakes, due to faulty and unreliable human judgement. Drone operators are physically removed from danger and the 'fog of war' is filtered and weakened through the drone's technological apparatus. And yet, the drones' promise of self-elevation above the frailty of human judgement and above the fog of war has proven illusory. Drone operators are still under pressure in making life and death decisions with limited knowledge and their humanity has allowed them to make mistakes, act recklessly, and target with prejudice. Nor are they themselves sufficiently distanced from the enemy and elevated above the consequences of their action. Studies have shown that drone operators suffer significant

---

[101]See International Human Rights and Conflict Resolution Clinic (Stanford Law School) and Global Justice Clinic (NYU School of Law), *Living Under Drones: Death, Injury, and Trauma to Civilians From US Drone Practices in Pakistan* (September 2012).

[102]On the practice of signature strikes see K. J. Heller, 'One Hell of a Killing Machine: Signature Strikes and International Law', (2013) *Journal of International Criminal Justice* 89.

[103]Such statements have been especially made in the context of drones. See J. Brennan (Assistant to the President for Homeland Security and Counterterrorism), 'The efficacy and ethics of US counterterrorism strategy', in J. Jaffer (ed.), *The Drone Memos* (2016), 199, at 207: 'it is hard to imagine a tool that can better minimize the risk to civilians'. This is also increasingly promised in the context of autonomous weapons. See the US submission at the 2018 GGE at www.unog.ch/ 80256EDD006B8954/(httpAssets)/7C177AE5BC10B588C125825F004B06BE/$file/CCW_GGE.1_2018_WP.4.pdf and the 'Remarks by Defense Department General Counsel Paul C. Ney Jr. on the Law of War', *Just Security*, 28 May 2019, available at www.justsecurity.org/64313/remarks-by-defense-dept-general-counsel-paul-c-ney-jr-on-the-law-of-war/.

[104]M. Schmitt, 'Autonomous Weapons Systems and International Humanitarian Law: A Reply to Critics', (2013) 4 *Harvard National Security Journal* 1; K. Anderson and M. Waxman, 'Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation Under International Law', in R. Brownsword et al. (eds.), *The Oxford Handbook of Law, Regulation and Technology* (2017), at 1097; see Arkin, supra note 5; J. Beard, 'The Principle of Proportionality in an Era of High Technology', in C. Ford and W. Williams (eds.), *Complex Battlespaces: The Law of Armed Conflict and the Dynamics of Modern Warfare* (2019).

[105]See, for example, the argument by Margulies that the superior pattern recognition capabilities of autonomous weapons will help in 'mapping affinities that can ripen into terrorist affiliations' and can therefore 'be immensely helpful in identifying previous unknown followers of ISIS or other groups and implementing a targeting plan'. See P. Margulies, 'Making autonomous weapons accountable: command responsibility for computer-guided lethal force in armed conflicts', in J. D. Ohlin (ed.), *Research Handbook on Remote Warfare* (2017), 405, at 422–3.

[106]I. S. Henderson et al., 'Remote and Autonomous Warfare Systems - Precautions in Attack and Individual Accountability', in Ohlin, ibid., at 335.

[107]See Professor Mary Cummings in Annex III of the 2018 GGE Report, para. 26: 'Due to the innate neuro-muscular lag of humans to perceive and act upon a situation . . . [LAWS] would be far more discriminatory provided existing computer perception issues were sorted out.'

[108]See M. Sassoli, 'Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Issues to be Clarified', (2014) 91 *International Law Studies* 308, at 320 (arguing that to the extent that autonomous weapons may be better at taking precautions commanders may have an obligation to use them).

[109]The US Department of Defense Artificial Intelligence Strategy, at 16, concurs: 'We will seek opportunities to use AI to enhance our implementation of the Law of War. AI systems can provide commanders more tools to protect non-combatants via increased situational awareness and enhanced decision support.'

post-traumatic stress.[110] And yet the promise persists: further, or full, automation will leave such flaws behind.

What is more, this trust in the self-elevating power of technology is inscribed in an overall perception of machines as more humane than humans,[111] not prone to sadism and bloodthirstiness, to panic and anger. It has been said that 'robots do not rape'.[112] That '[t]hey can be designed without emotions that cloud their judgment or result in anger or frustration with ongoing battlefield events'.[113] That they are free from the 'fear and hysteria' that push humans towards 'fearful measures and criminal behavior'.[114]

Technology and automation may, therefore, elevate us above our inhumane humanity. Technology will thus limit the circumstances under which we need to revert to our faulty judgement, especially under pressure. Human judgement, in the conduct of war, is perceived as weak and unreliable. Distance from it is seen as improving, even salutary. We delegate judgement to machines and the de-humanization of war entails its civilization.

When President Kennedy's scientific advisor referred to the victory in the 'arms race against ourselves', in the quote opening this section, he was not making a philosophical point. He was, rather, referring to the US overcoming difficulties developing the weapons technology that would allow it to win the Cold War. And yet, the arms race against ourselves can be understood to reflect a more fundamental struggle: against human weakness, in body and mind; against the imperfection of our humanity. Technology and war, technology *in* war, are the battlefronts and law has a part to play.

### 3.3. The inherent compatibility of legal technics

Elevation above oneself relies on mechanical rationality. The distance it creates from the frailty of human judgement stands on the shoulders of an increasingly hardwired way of thinking. Thinking as calculation, and the perception of the human mind as a machine, have a long tradition in the trajectory of rationalist philosophy. From Hobbes[115] through Leibniz[116] to Descartes[117] the metaphor of mind as machine inspired both a rich philosophy of science and an intense urge for

[110]See W. Chappelle et al., 'An analysis of post-traumatic stress symptoms in United States Air Force drone operators', (2014) 28(5) *Journal of Anxiety Disorders* 480; Chamayou, *supra* note 78, at 117–19.

[111]See Arkin, *supra* note 5, at xvi. See also Chamayou, ibid., at 208–9 on the distinction between ontological and axiological humanity.

[112]See this position in the otherwise critical report of Special Rapporteur Heyns, at para. 54: '[Lethal Autonomous Robots] will not be susceptible to some of the human shortcomings that may undermine the protection of life. Typically, they would not act out of revenge, panic, anger, spite, prejudice or fear. Moreover, unless specifically programmed to do so, robots would not cause intentional suffering on civilian populations, for example through torture. Robots also do not rape.' See K. Bergtora Sandvik and K. Lohne, 'Lethal Autonomous Weapons: Killing the "Robots-don't-Rape" Argument', *IntLawGrrls Blog*, 5 August 2015, available at ilg2.org/2015/08/05/lethal-autonomous-weapons-killing-the-robots-dont-rape-argument/; J. Turner, 'We should regulate, not ban killer robots', *Spectator*, 28 August 2017, available at blogs.spectator.co.uk/2017/08/we-should-regulate-not-ban-killer-robots/.

[113]See Arkin, *supra* note 5, at 29.

[114]Ibid., with reference to M. Walzer, *Just and Unjust Wars*.

[115]See S. Franchi and G. Güzeldere, 'Machinations of the Mind: Cybernetics and Artificial Intelligence from Automata to Cyborgs', in Franchi and Güzeldere, *Mechanical Bodies, Computational Minds: Artificial Intelligence from Automata to Cyborgs* (2005), 15, 40–1.

[116]See P. Beeley, 'Leibniz and Hobbes', in B. Look (ed.), *The Bloomsbury Companion to Leibniz* (2014), 32–51. See also the influence of Leibniz on Charles Babbage and the creation of his Analytical Engine, the first generally programmable machine, in S. Bullock, 'Charles Babbage and the Emergence of Automated Reason', in P. Husbands et al. (eds.), *The Mechanical Mind in History* (2008), 19–40.

[117]While Descartes was not, as a whole, a materialist and believed that reason was beyond the reach of mere machines his role in mechanistic thinking is explored in M. Wheeler, 'God's Machines: Descartes and the Mechanization of Mind', in Husbands et al., ibid., at 307.

scientism.[118] Descartes' ambitious parallels between machines and non-human animals and his speculation on the creation of indistinguishable automata place him at the centre of this tradition, even though his clear separation of mind from body allowed his cognitive philosophy to remain free of his materialism. To the extent, however, that the mind and cognition are identified with the brain, Cartesianism may 'degenerate' into machinist disembodied cognition.[119] Alongside the submission of the mind, through the brain, to mechanical description, a strand of logical positivism centres on the symbolic representation of the world. For example, in the twentieth century, Gottlieb Frege 'showed that rules could be formalized so that they could be manipulated without intuition or interpretation'.[120]

Elizabeth Boden has documented in detail how this tradition of 'mind as machine' and formal logic are central in the development of both the ambitions and the philosophy of cognitive science and artificial intelligence.[121] The rich history of this relationship and the intricacies of the philosophical debate are beyond the scope of our inquiry. And yet that understanding of mind as machine and of thinking as, however complicated, symbolic representation can be found in certain strands of legal thinking, influencing the role of legal scholarship and practice on technology in war. Indeed, the process of and the quest for the establishment of principles, precedents, predictable outcomes and the overall professionalization and formalization of judgement (in warfighting) has been central to the contemporary law of armed conflict project.

There are important parallels between the aspirations and hopes assigned to technology and those assigned to law. Chief technological optimists today, indeed, suggest that automation will simply play the role of making sure the law is enforced,[122] to the extent that 'bad apples' and the 'fog of war' do not interfere. This attitude is especially strong in non-lawyers, perhaps prone to simplistic versions of the law,[123] but it also finds fertile ground in different strands of mainstream legal analysis.

The dominant discourse approaches the present and future regulation of autonomous weapons as a question of compatibility. What is asserted is a set of rules and what is required is for a machine – both the machine's hardware and its software – to be able to meet and implement these rules. The question becomes, for example, whether the principle of distinction or the avowedly more complex principle of proportionality[124] can be articulated in a series of logical steps and whether the machine's technological capacity, for example in image recognition of the received characteristics of a civilian target, is able to perform such a code. Such logical steps can be set out as a mathematical formula[125] or as programming language.[126] Myriad questions can be posed about the specifics of such categorization, but I am concerned with the overall stance. And this is one of wait-and-see. The position could be simplistically stated thus:

---

[118]See R. Williams, 'Introduction', in R. Williams and D. Robinson (eds.), *Scientism: The New Orthodoxy* (2015); M. Boudry and M. Pigliucci (eds.), *Science Unlimited? The Challenges of Scientism* (2017).

[119]This is also at the centre of Noll's discussion of neurotechnology. See *supra* note 14, at 219–23 with some reference to the literature critiquing 'degenerate Cartesianism' in the philosophy of neuroscience.

[120]H. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason* (1992), xi.

[121]E. Boden, *Mind as Machine: A History of Cognitive Science*, vols. I & II (2006).

[122]See Arkin, *supra* note 5.

[123]Conversely, idealized versions of the law often feed blanket opposition to autonomous weapons systems, as will be discussed further below.

[124]M. Wagner, 'The Dehumanization of International Humanitarian Law: Legal, Ethical and Political Implications of Autonomous Weapons Systems', (2014) 47 *Vanderbilt Journal of International Law* 1371, at 1393.

[125]For a recent attempt at articulating the basic principles of IHL as a set of mathematical formulae see M. Schmitt and Major M. Schauss, 'Uncertainty in the Law of Targeting: Towards a Cognitive Framework', (2019) *Harvard National Security Journal* 148. It should be noted the authors are, however, careful to point out, at 152, that '[t]he formulae should not be viewed as an attempt to reduce targeting decisions to mechanical deterministic calculations' and that their formulations aim to inform the choices of operators, rather than to be imposed through an algorithm.

[126]See Arkin, *supra* note 5, and, for example, his programming algorithm for the principle of proportionality at 186.

Currently we do not have the technology that would perform a distinction or proportionality calculation. While we don't have answers to all of the challenges, maybe we will in the future. In which case, the autonomous system will go through a weapons review and this will determine its compatibility with the law of armed conflict.[127]

One constituency taking such an approach are 'pragmatist' (military) lawyers, who, while of course cognizant of legal complexities, especially those associated with notoriously difficult to apply principles such as proportionality, are open to the logic of the law's codification in algorithms. Such a task of codification is essentially seen as a question for engineering. When the technology will allow it, there is no reason why codification may not occur.[128] Therefore, even if one does not rush to diagnose legal salvation through technology *yet*, one does not see why law and automated decision making would not be compatible.

Indeed, jurisprudential approaches especially associated with a strand of philosophy of logic,[129] arguably go further, seeing law and technology as inherently compatible and artificial intelligence as an ideal avenue to discuss questions of legal logic and the categorization, interpretation and application of rules.[130] Law itself is seen as potentially profiting from the tools of formal logic associated with artificial intelligence, the two constituting a mutually improving and reinforcing relationship of formalization.[131] There are two things needed: 'contents to be inserted in the knowledge base, and the choice of formalism (with related formal inferential procedure) in which to represent those contents'.[132] Judgement is, again, reserved on the extent to which technology is currently able to provide the tools for legal interpretation and application – for example in matters of visual recognition, natural language processing, adaptability. And the especial difficulties posed by contextual and qualitative judgements in the application of certain legal rules is recognized.[133] The position about the future is agnostic; or, it is in abeyance. The answer remains to be seen, complex problems of engineering, beyond current science, will need to be resolved. Of course, 'the perfectibility of man is absolutely indefinite'.[134] And law is treated as inherently technological; law *is* a technology – our self-perfection and civilization will occur through law and technology in tandem.

This is not to discount the potential rigor of analytical logic, the usefulness of some computational tools[135] or the use of pragmatist professionalism when encountering issues in the law's application, including in the context of new weapons technology. It is, however, to suggest that both stances may display a tendency to uncritically embrace a reductive approach to law, through

---

[127]The expert contribution of W. Boothby at the 2015 CCW Meeting of Experts is typical: 'We do not know whether future technology may produce weapon systems that can out-perform humans in protecting civilians and civilian objects. It would in my view be a mistake to try to ban a technology on the basis of its current shortcomings, when in future it may actually enable the law to be complied with more reliably than now.', available at www.unog.ch/80256EDD006B8954/(httpAssets)/616D2401231649FDC1257E290047354D/$file/2015_LAWS_MX_BoothbyS+Corr.pdf.

[128]See R. Williams, 'Introduction', in R. Williams and D. Robinson, *Scientism: The New Orthodoxy* (2015) 1, at 2: 'Technology is the ultimate pragmatism, and in a world dominated by technology, and intoxicated by technological solutions to practical problems, everything is viewed as "standing in reserve," ready to be used by or subjected to some sort of technology or technological process.'

[129]See generally the work of G. Sartor and the editors and contributors in *Artificial Intelligence and Law*.

[130]See G. Sartor, 'Artificial Intelligence in Law and Legal Theory', (1992) 10 *Current Legal Theory* 1.

[131]Ibid., at 2: 'it is true that the analyses of legal theory have rarely attained the level of specificity and precision required by computability, especially because the use of formal methods in law has been limited to a few exceptions. Nevertheless, even informal analyses can be of enormous importance for AI, which is exposed to simplifications and reductionisms also because of its need for formalization'.

[132]Ibid., at 23.

[133]Ibid., at 36.

[134]Condorcet, *Sketch for a Historical Picture of the Progress of the Human Mind* (1795).

[135]See W. Alschner, 'The Computational Analysis of International Law', in R. Deplano and N. Tsagourias (eds.), *Research Methods in International Law: A Handbook* (2019), for the developing techniques in the 'mining' of 'international law as data'.

technology, one that will not do justice to its substance, one that pursues the dehumanization of judgement in the service of double elevation. Law as technology, as formal logic to be engineered in artificially intelligent machines, can be seen as rising above the frailty of human judgement.

## 4. Against Double Elevation

### 4.1. Angst

The trajectory of optimism has always gone hand in hand with angst. Indeed, at the very start of post-war futuristic engagement optimism and pessimism coexisted. While Norbert Wiener preached the coexistence and self-regulating adaptation of human/machine in his book *Cybernetics*, his intellectual integrity allowed him to repeatedly deplore the social, political and moral dangers of automation, both in a companion volume written for the general public[136] and in his interactions with increasingly starry-eyed disciples.[137]

The prosecution of war heightens such angst. Increasingly, the perils of double elevation are recognized. The trajectory of the debates over the use of 'smart bombs' in Kosovo[138] or Iraq to the drone-enabled 'War on Terror' evokes the distancing that elevation above the enemy may produce over the often de-humanized other and the production of indefinite asymmetrical global war. Similarly, elevation above oneself has been perceived to lead to 'the fabrication of political automata'[139] and the loss of freedom. The latter point is meaningfully set out in the context of armed drones by Roger Berkowitz, director of the Hannah Arendt Centre:

> In the end, the threat drones pose is not only to civilians in war or to jobs. The real threat is that as our lives are increasingly habituated to the thoughtless automatism of drone behavior, we humans habituate ourselves to acting in mechanical, algorithmic, and logical ways. The danger drones pose, in other words, is the loss of freedom.[140]

Beyond double elevation, this angst over the loss of judgement, over the loss of control over the moral parameters of war-fighting and decision-making feed into the wider existential fear associated with technological pessimism, namely that the emancipation of the creation will be complete, lost to its creator. This concern, at the level of prediction, fears that we are nearing a 'singularity' where artificial intelligence will fully escape human control. While some have hailed this coming singularity in near-religious terms,[141] and some, associated with the transhumanist movement, invest in what they perceive as our overcoming our essential human weakness and even overcoming death,[142] others are pausing in existential dread.[143] While it is important to separate the angst associated with the loss of control over specific tasks from that over a wider 'sorcerer's apprentice' deluge, such fears are as interrelated as the aspirations that feed them.

What is to be done? How do those not sharing in the enthusiasm of scientism or an agnosticism of piece-meal problem solving engage in the present formation of the future of war, technology, and law?

---

[136]See N. Wiener, *The Human Use of Human Beings: Cybernetics and Society* (1950); N. Wiener, 'Some moral and technical consequences of automation', (1960) 131(3410) *Science* 1355. This is developed in more philosophical depth in his *God & Golem, Inc.: A Comment on Certain Points where Cybernetics Impinges on Religion* (1964).

[137]One of these was Alice Mary Hilton whose *Logic, Computing Machines and Automation* (1963) placed its hopes on automation for 'human beings [to] become truly civilised'. See Wiener's letter of 8 March 1963 in Rid, *supra* note 90, at 103.

[138]See P. Kahn, 'The Paradox of Riskless Warfare', (2002) *Philosophy & Public Policy Quarterly* 2.

[139]See Chamayou, *supra* note 78, at 205.

[140]R. Berkowitz, 'Drones and the Question of the "Human"', (2014) 28(2) *Ethics & International Affairs* 159, at 169.

[141]See Kurzweil, *supra* note 2.

[142]See O'Connell, *supra* note 2 for interviews with current adherents of the faith.

[143]For example, Bostrom, *supra* note 4. Importantly Bostrom was a member of the transhumanist movement; his book has been recently recommended by techno-optimists such as Elon Musk and Bill Gates; his Oxford Project is funded, partly, by the former.

The major stance in opposition to technological escalation towards full autonomy is centring on the taboo of delegating life/death decisions to machines and the innate inability of machines to properly apply law. A public expression of this position is that of the Campaign to Ban Killer Robots.[144] The campaign is notable in resisting both superficial techno-optimism and dangerously instrumental pragmatism. The potential power of mobilization of public opposition notwithstanding, the position presents some analytical, strategic and conceptual shortcomings.

Firstly, the Campaign's approach starts from an assumption of a fundamental 'change of paradigm'.[145] This does not always appreciate the continuing role of technology in war and its function in the double elevation described here. Indeed, I have argued that the escalation towards full autonomy and merged heteronomy represents a continuation of an existing trajectory of distancing through the elevation above one's enemy and above oneself, albeit with a perhaps justified presentiment of an 'avalanche', a violent acceleration of pace, out of control.

Secondly, present practice does not suggest that a prospect of successful imposition of a ban or moratorium is realistic. This is due both to the anticipated military advantage and the inscription of this process in socioeconomic structures and expectations, as reflected in the evidence of enthusiastic investment in the acceleration of this trajectory. Instead, as the analysis above has suggested, while the absolute of full autonomy (combining kinetic and cognitive elements) is kept at bay, the ground is constantly prepared.

Thirdly, the primary focus on preventing full autonomy or, inevitably anthropomorphized, 'killer robots' is in danger of missing the target. Autonomy is complemented by increasingly merged heteronomy. As discussed in Section 2.1 above, merged heteronomy addresses the logistical limitations of spatially spread human/machine networks. Crucially, while presenting itself as respecting the moral taboo of life-and-death delegation, merged heteronomy advances the mechanization of judgement in pursuit of double elevation. To the extent that the maintenance of 'meaningful human control' is primarily focused on 'keeping humans in the loop', it is in danger of ignoring the gradual change in the nature and function of that very loop. As reliance on artificial intelligence increases, it is humans who are becoming the 'killer robots'.

Finally, to the extent that the position relies on the incompatibility of autonomous weapons with international humanitarian law, it may be vulnerable to the complicity, discussed in Section 3.3, of certain strands of legal thinking with an understanding of knowledge as 'a large store of neutral data'[146] and the promise of the piece-meal resolution of technical legal problems. It also allows one's intuitive angst to be assuaged by promises of, or indeed steps towards, the panacea of global regulation.[147] As important as such regulation may be in structuring the ambitions of both state and private actors, it would not, *per se*, address the most fundamental dangers of the mechanization of judgement.

Law will neither ban nor regulate away what causes our angst. To the contrary, it may be adapted to serve mechanized judgement. If we are to oppose double elevation and the mechanization of judgement, and hope to use law to this effect, we need legal thinking to serve this purpose. Otherwise, all we can do is surrender to the stance of agnostic abeyance, until the code is engineered.

---

[144]See www.stopkillerrobots.org; see also Human Rights Watch, 'Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban', 9 December 2016, available at www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban.

[145]J. Kellenberger, 'International humanitarian law and new weapon technologies', *International Institute of Humanitarian Law*, September 2011, available at www.unog.ch/80256EDD006B8954/(httpAssets)/F77AF42ED509F890C1257CD90025183F/$file/IHL+&+new+weapon+technologies_Sanremo.pdf.

[146]H. Dreyfus, *What Computers Can't Do* (1972), at 190.

[147]Indeed, the recently released first US Department of Defense Artificial Intelligence Strategy Summary gestures, at 15, towards 'a global set of military AI guidelines'.

## 4.2. Irreducible intelligence, and irreducible law

An opposition to the present future of the loss of judgement requires an understanding of law as irreducible. I have argued that the evolution of new weapons technology towards increasing autonomy and merged heteronomy serves, and accelerates, a double elevation, above one's enemy and above oneself, which pursues the mechanization and distancing of judgement; that to the extent that the role of law, in this context, is viewed as a question of 'compatibility' or 'adaptability', there is a danger that it, too, would serve this purpose. And yet, while law is no panacea to be administered through regulation or outright proscription, it does not have to be the handmaiden of mechanization. In this section, I conclude by arguing that to think of the law of war in a way that resists the demands of double elevation, we should turn to the philosophical and sociological critique of the cognitive science that buttresses much of the existing logic of artificial intelligence. Alongside our understanding of the historical and material process whereby double elevation and increasing autonomy are produced, outlined above, the critiques of the epistemology of artificial intelligence are a necessary guide for the appreciation, defense, and practice of irreducible legal thought.

Gregor Noll, in his analysis of the influence that the weaponization of neurotechnology has on international humanitarian law (IHL), uses the critiques of positivist cognitive science and its reliance on antiquated or simplistic understandings of cognition 'reducing high-level behaviors to low-level, mechanical explanations, formalizing them through pure scientific rationality'.[148] He highlights what he calls the 'degenerate cartesianism'[149] of neuroscience, which is expressed in the separation between perception and cognition. Isolating cognition, including legal cognition, from human perception, he argues, reduces 'the legal knowledge of the IHL experts … to a set of skills regarding a particular procedure to be followed in decision-making and a range of outcomes'.[150]

To understand this process of reduction, the intellectual history of artificial intelligence and the fierce debates around its epistemological qualities are instructive.[151] A 'representationist' theory of mind, understanding thought as the computation of symbols representing reality,[152] trusted and employed, with varying sophistication, by generations of artificial intelligence research, has been consistently criticized since the beginnings of 'good old-fashioned AI (GOFAI)'[153] by Hubert Dreyfus. A philosopher infiltrating the MIT AI community, Dreyfus's 1965 RAND Corporation memorandum[154] and his 1972 book[155] constituted a full-frontal attack on an emerging and ambitious project. Dreyfus argued that representational thinking fails to account for the:

> know-how, [which,] along with all the interests, feelings, motivations, and bodily capacities that go to make a human being, would have had to be conveyed to the computer as knowledge. … [M]aking our inarticulate, preconceptual background understanding of what it is like to be a human being explicit in a symbolic representation [is] a hopeless task.[156]

---

[148]P. Dourish, *Where the Action Is: The Foundations of Embodied Interaction* (2004), at vii.

[149]See Noll, *supra* note 14, at 220–3.

[150]Ibid., at 223

[151]These were no bloodless academic affair. The personal nature of the debate arguably reflected fundamental differences in respective epistemologies and can be seen in the introductions to the different editions of Dreyfus's *What Computers (Still) Can't Do*, the reviews of the book's re-issue in the special issue of vol. 80 (1996) of *Artificial Intelligence*, and Boden *supra* 121, Ch. 11.

[152]For a discussion of different, 'strong and weak', representationist theories see F. Egan, 'Representationism', in E. Margolis et al., *The Oxford Handbook of Philosophy of Cognitive Science* (2012), 250.

[153]The term denotes the first phase of artificial intelligence, from the 1950s to the 1990s, also referred to as 'symbolic AI'. It was coined by J. Haugeland in *Artificial Intelligence: The Very Idea* (1985).

[154]H. Dreyfus, *Alchemy and Artificial Intelligence* (1965), available at www.rand.org/pubs/papers/P3244.html.

[155]See Dreyfus, *supra* note 120.

[156]Ibid., at xvi–xvii.

Symbolic representation, Dreyfus insisted, can only take you so far. According to this view, the artificial intelligence projects, from the 1950s to the 1990s, attempting to input a map of reality in a machine, sometimes through the use of distinct 'micro-worlds' which would be used for generalizations, were pointless and doomed to fail.[157] Outside the battles of academia, the futility of perfect representation is, perhaps, best conveyed in a story by Jorge Luis Borges, entitled *On Exactitude in Science*, so short it may be cited in its entirety:

> In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast Map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.[158]

A perfect code is impossible, futile and self-defeating. It is not only that we may not be able to agree on and articulate the primary rules to be input in the system,[159] it is that once they are abstracted from their human masters they become 'counterfeit'. This applies to law as much as it applies to the rules of cognition in general.[160]

 Machine learning neural networks do not overcome this problem. While these, unlike symbolic artificial intelligence, do not simply reproduce a symbolic representation map, but develop 'a history of input-output pairs',[161] their learning and adaptive behaviour is second order, as it follows pre-set parameters. While quantitively sophisticated, they mimic rather than think.[162] Even if mapping and generalization are adaptive and self-generating, the counterfeit problem remains. 'Deep learning' is as vulnerable to the critique as symbolic representation was. It is not only that artificial intelligence has, so far, failed beyond the reproduction of mechanic tasks. It is that it can never succeed. In his final piece of writing, more than 40 years after his first missive, Dreyfus argues that, because 'how we directly pick up significance and improve our sensitivity to relevance depends on our responding to what is significant for us' given our particular characteristics, for a non-representational artificial intelligence to be successful:

> we would not only need a model of the brain functioning underlying coupled coping . . . , but we would also need—and here's the rub—a model of our particular way of being embedded and embodied such that what we experience is significant for us in the particular way that it is. That is, we would have to include in our program a model of a body very much like ours with our needs, desires, pleasures, pains, ways of moving, cultural background, etc.[163]

---

[157]Ibid., at 4–27.

[158]J. L. Borges, *The Aleph and Other Stories* (1945), at 181.

[159]See, for example, A. Haque's point that the disagreements on proportionality 'are substantive, not semantic, and rooted in its contestability, not its vagueness', in 'Indeterminacy and the Law of Armed Conflict', (2019) 95 *International Law Studies* 15.

[160]One of the first cognitive scientists to appreciate Dreyfus's critique recognizes that phenomenologists 'point out that in such an approach we are no longer using the theoretical terms in their original authentic sense: We have abandoned the original problem and find ourselves talking about a counterfeit world in place of the one we initially intended to study – a formal world of logic, mathematics and "bits" of information in place of a human world of experience, knowledge and purpose'. See Z. Pylyshyn, 'Minds, machines and phenomenology: Some reflections on Dreyfus' "What computers can't do"', (1974) 3(1) *Cognition* 57, 63.

[161]Dreyfus, *supra* note 120, at xv.

[162]See also Berkowitz, *supra* note 140, at 165, on the question of drone-generated art.

[163]H. Dreyfus, 'Why Heideggerian AI failed and why fixing it would require making it more Heideggerian (2007)', in H. Dreyfus and M. Wrathall, *Skillful Coping: Essays on the phenomenology of everyday perception and action* (2014), 250, 272–3.

Our distinct cognitive functions cannot be abstracted from our overall existence. Importantly, while expressed in the language of functionality and prediction, Dreyfus's discussion of 'what computers *can't* do' entails what they *should not* (try to) do. Using the emotive language of 'obscenity' and 'disgust', Joseph Weizenbaum's[164] epistemology is more avowedly existentialist: 'an organism is defined, in large part, by the problems it faces. Man faces problems no machine could possibly be made to face. Man is not a machine'.[165] Indeed, to the question posed to him by the eminent artificial intelligence researcher John McCarthy 'What do judges know that we cannot tell a computer?'[166] Weizenbaum responded violently: 'The very asking of the question … is a monstrous obscenity. That it has to be put into print at all, even for the purpose of exposing its morbidity, is a sign of the madness of our times.'[167]

If Dreyfus and Weizenbaum decry the futility and impoverishment of cognitive reduction, the sociology of science highlights the consequences in human/machine systems. One of the reviewers of the 1992 re-issue of Dreyfus's book, Harry Collins, argued that Dreyfus did not go far enough in appreciating the social embeddedness of both computers and humans as well as, crucially, the concepts that humans are using.[168] There was no independently stable knowledge to be input in the computer in the first place. This unstable knowledge, reduced through programming, is then fed back to humans with the veneer of disembodied objectivity. Twenty-two years later, it is the poverty rather than the dreaded omniscience of artificial intelligence which is the major threat:

> As it is, the big danger facing us is not the Singularity; it is failing to notice computers' deficiencies when it comes to appreciating social context and treating all consequent mistakes as our fault. Thus, much worse, and much more pressing than the danger of being enslaved by enormously intelligent computers, is our allowing ourselves to become the slaves of *stupid* computers – computers that we take to have resolved the difficult problems but that, in reality, haven't resolved them at all: the danger is not the Singularity but the Surrender![169]

The critiques of the epistemology and sociology of artificial intelligence must inform our understanding of its relationship with law. They can help us understand what is lost and impoverished when legal concepts and rules are reduced to algorithm – the futility, but also the harm in the pretension of perfect representation and reproduction. They can also help us understand the process of both intellectual and moral impoverishment in the removal of intelligence from its social context and the outsourcing, through formalization, of life-and-death decisions to mechanized judgement. The 'surrender' Harry Collins refers to above, can be understood in the context of law as an abdication of responsibility. I am not referring here, *stricto sensu*, to the stretching of individual liability to the breaking point in complex human/machine systems, but to the responsibility inherent in our interactions with the law. Legal rules and principles, such as that of proportionality, reflect a combination of values that are both meaningful and problematic. It is our task, through our situated moral intelligence, flawed as it is, to take responsibility for the rules governing our violence and to act as custodians, interpreters, and appliers of the law. All the more

---

[164]See especially J. Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation* (1976).

[165]Ibid., at 203.

[166]McCarthy's answer was 'nothing', as he thought 'all human knowledge could be formally represented by AI'. See Boden, *supra* note 121, at 851. For the work of McCarthy, 'father of AI', according to his own website, see jmc.stanford.edu/. McCarthy was also one of the least patient reviewers of Dreyfus's work. See J. McCarthy, 'Review of Dreyfus, *What Computers Still Can't Do*', (1996) 80 *Artificial Intelligence* 143.

[167]See Weizenbaum, *supra* note 164, 226–7.

[168]See H. Collins, 'Embedded or embodied? a review of Dreyfus, What Computers Still Can't Do', (1996) 80 *Artificial Intelligence* 99, 101–5 referring to late Wittgenstein.

[169]H. Collins, *Artifictional Intelligence* (2018), 5.

so when these rules and principles are applied by human beings who risk their lives for the collective against the lives of other human beings, in other collectives.

If there is a task for law in countering double elevation, it is not discharged through regulation, but it requires the defense of its complexity, its subtlety and its humanity in a way that resists a mechanistic philosophy of cognition. Legal thinking should not avoid the uncertainty and incommensurability of situated judgement but embrace it. This is judgement that is situated in human relationships within collective structures; judgement in the immediacy of decision-making in battle, which cannot be conclusively determined *a priori*; judgement when real human beings are making impossible choices.

The quest for the exercise of irreducible intelligence in law does not entail a distinct jurisprudential preference. While the appreciation of the complexity and indeterminacy of legal meaning and the decisive role of social context in both the making and application of the law may be associated with the anti-formalism of pragmatism, critical legal studies or the third world approaches to international law, this is not the goal of this article. Despite the fact that certain strands of formalist or mainstream lawyering seem to be comfortable with their reducibility to algorithmic input, legal thinking and judgement ought to display a situated intelligence irrespective of their jurisprudential politics. Indeed, the law of war will be applied in battle, by soldiers who are unlikely to have subscribed to a particular school of jurisprudence.

Doctrinal scholarship is perfectly capable of irreducible subtlety. Indeed, the challenge that automation poses to positivist scholarship is not the formulation of rules in a way that can be coded, but the contribution to the clarity of the law in a way that cannot be automatically reproduced. But it may also be that the international law of war, in order to address the increasing pressures of double elevation, should step back to question, historicize, and theorize[170] its fundamental concepts and their application. The scholars and practitioners of the law of war should study the evolving sociotechnical landscape, and our gradual immersion in human/machine systems.[171] They should take on the epistemological challenge of engaging the elements of its situated subjectivity – the discretion,[172] emotion,[173] imagination,[174] passion[175] in the rules and their application.

## 5. Conclusion

The value of the law is precisely in what cannot be grasped in computational form and our legal thinking should reflect this. The present future of the mechanization of judgement should not be seen as a radical break with the past, an alien invasion. It is, rather, the accelerated evolution of an impoverished tradition. Both increasing automation and merged heteronomy are advancing the distancing and mechanization of judgement in the service of double elevation – against our enemies and against our very selves. While we fear losing our place 'in the loop', that very loop is

---

[170]On the relative under-theorization of the laws of war because of the pro-active and pragmatic nature of humanitarianism see F. Mégret, 'Theorizing the Laws of War', in A. Orford and F. Hoffmann (eds.), *The Oxford Handbook of the Theory of International Law* (2016), 762.

[171]See, in this direction and using cybernetic theory, L. Suchman and J. Weber, 'Human-Machine Autonomies', in Bhuta et al., *supra* note 16, at 72. See also H.-Y. Liu, 'From the Autonomy Framework towards Networks and Systems Approaches for "Autonomous" Weapons Systems', (2019) 10 *Journal of International Humanitarian Legal Studies* 89.

[172]On discretion focusing on autonomous weapons see E. Lieblich and E. Benvenisti, 'The obligation to exercise discretion in warfare: why autonomous weapons systems are unlawful', in Bhuta et al., *ibid.*, at 245.

[173]See, for a recent example, M. Deland et al. (eds.), *International Humanitarian Law and Justice: Historical and Sociological Perspectives* (2019), especially R. Sutton, 'A Hidden Fault Line: How International Actors Engage with International Law's Principle of Distinction'.

[174]For an essay-length introduction focusing on judging see M. Del Mar, 'The Legal Imagination', 2017, *Aeon*, available at aeon.co/essays/why-judges-and-lawyers-need-imagination-as-much-as-rationality.

[175]N. Modirzadeh, 'Cut These Words: Passion and International Law of War Scholarship', (2019) *Harvard Journal of International Law* (forthcoming).

changing and, potentially with the complicity of legal technics, we are in danger of becoming the killer robots we want to ban. This article aims to contribute to the recognition of these dangers and to begin articulating a response. It is a call for legal thinking to discharge its disciplinary role within a broader philosophical and political battle. If this direction of thought is not prioritized now, the powerful forces of double elevation may reduce law to something that is both unrecognizable and all too recognizable.