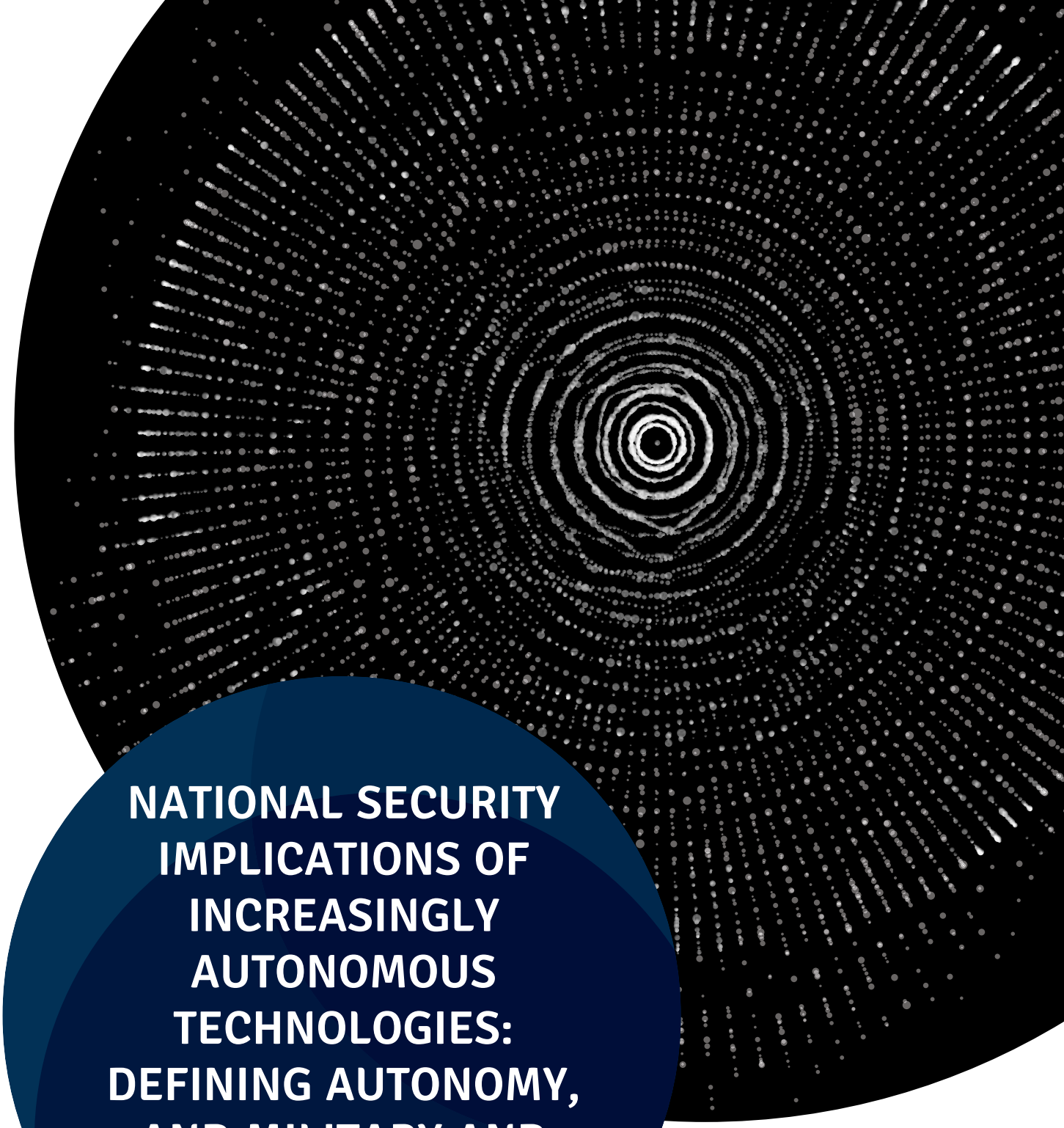


NATIONAL SECURITY IMPLICATIONS OF INCREASINGLY AUTONOMOUS TECHNOLOGIES: DEFINING AUTONOMY, AND MILITARY AND CYBER-RELATED IMPLICATIONS

PART 1

Policy Report

February 2015

Caitríona H. Heinl

RSiS
Nanyang Technological University
S. RAJARATNAM SCHOOL OF INTERNATIONAL STUDIES

Policy Report

NATIONAL SECURITY IMPLICATIONS OF INCREASINGLY AUTONOMOUS TECHNOLOGIES: DEFINING AUTONOMY, MILITARY AND CYBER-RELATED IMPLICATIONS

Part 1

Caitríona H. Heinl
February 2015

Centre of Excellence for National Security (CENS),
S. Rajaratnam School of International Studies (RSIS),
Nanyang Technological University

Executive Summary

This is the first of a two-part report that highlights the mounting importance for the national security agenda of technologies that are becoming increasingly autonomous, or becoming gradually more independent of human control in other words. At present, it is still relatively unclear how maturing autonomous technologies, including potentially fully autonomous and lethal systems, might impact national security exactly in terms of military and economic implications, or possible misuse by criminals. This two-part report finds that many questions still remain unaddressed and that there are several significant policy gaps that should be further analysed.

While some aspects of this area are still in their infancy, the full report aims to identify the key questions that are beginning to emerge. It also highlights the salient aspects of several discussions that have been recently initiated and will impact national security. Thus far, as a United Nations Institute for Disarmament Research (UNIDIR) report of March 2014 notes, there has been a lack of critical analysis on how the proliferation of increasingly autonomous systems might alter regional security dynamics.[1] China, for instance, recently became the largest buyer of industrial robots, overtaking Japan for the first time with an approximately 60 per cent increase in a one-year period from 2012 to 2013.[2] And while scientists, ethicists, and futurists, amongst others, have hotly debated several gaps marked within the report in the past, wider policy circles are only recently beginning to seriously consider these questions to the same extent. This two-part report argues that these issues now require deeper consideration and it is an opportune time to shape the strategic debate.

The United Nations Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns, recently explained that while the technology for drones is already in use and discussions are now being held on their regulation, autonomous robotics presents a unique situation since the technology is not actively used yet.[3] This therefore presents some unique challenges, which are addressed throughout both parts of the report. The opening section of this first part of the report discusses the nature of maturing autonomous technologies and the significance of potential lethality. It finds that, although there is an increasing military interest in this area, a clear understanding of the nature of these technologies is still lacking in the policy community.

The next section then provides an outline of several broader military implications as well as cyber-related implications that could arise in this area. It is likely that states will pursue technological superiority via increasingly autonomous technologies for both economic and military reasons. Yet, deeper analysis of the long-term implications is needed in terms of possible military advantages and disadvantages that might ensue, including the role of the human vis-à-vis the machine, as well as how military interest in autonomy might evolve globally. Given geopolitical uncertainties in the Asia Pacific region, such developments could also be significant if states seek technological superiority with autonomous technologies.

The second part of this report analyses the challenges of controlling and regulating this space. While various stakeholders have made numerous recommendations, there does not

---

1 UNIDIR Resources, "Framing Discussions on the Weaponization of Increasingly Autonomous Technologies", March 2014, 8.

2 Tanya Powley, "China becomes largest buyer of industrial robots", http://www.ft.com/cms/s/0/a5cca8c0-e70c-11e3-aa93-00144feabdc0.html#axzz35X1ZGoLX, 1 June 2014.

3 Christof Heyns, Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, "Lethal Autonomous Robotics", United Nations Institute for Disarmament Research (UNIDIR) Conference, http://www.unidir.org/programmes/security-and-society/lethal-autonomous-robotics, 23 May 2013.

seem to be a silver bullet solution at this juncture. Moreover, the report finds that there are several highly significant legal ambiguities, which require clarification.

Furthermore, these technologies often have a dual-use nature – for both military application and civilian purposes, and both the public and private sectors are driving these developments by investing heavily in R&D in pursuit of their own objectives. This part of the report finds that while innovation and economic growth should not be disproportionately stifled, stronger collaboration between the public sector and industry, as well as academic research laboratories, is advisable to shape policies responsibly and manage unexpected developments that could perhaps be detrimental. Malicious non-state actors also add a further layer of complexity since terrorist groups, organised crime gangs, or proxy actors could possibly obtain or alter commercially available technologies.

The last section of the second part of the report finds that the ethical implications of these tools require deeper consideration, and public perception of such advanced technologies is another important factor that should be considered.

Policy Uncertainty: Challenges and Opportunities

From Automation to Increasingly Autonomous Technologies

According to several recent analyses, increasing autonomy[4] in machines and systems such as robots, weapons, and weapon systems is being driven by advances in robotics, machine learning, artificial intelligence (AI), computational power, networking, engineering, and other disciplines.[5] While it seems that states are not using fully autonomous robots yet, UN Special Rapporteur Christof Heyns explains that the technology appears to be available or is at least becoming available very soon. He therefore considers that lethal autonomous robotics is the next generation of weaponised technology after drones.[6] While such technological developments could have positive consequences, like past inventions, including unintended results, some developments could also lead to threats.[7] This section therefore outlines the nature of such autonomous technologies and explains why they are growing in significance in the national security domain.

Increasingly autonomous technologies are defined in this context as technologies that are becoming increasingly independent of human control, albeit to varying degrees. So far, between 50-80 countries are developing robots and/or have made operational use of robots in the battlefield. In South Korea and Israel for example, robotic sentries with the capacity to be armed have already been deployed.[8] U.K. government reports further assert that states will focus investments on developing capabilities and countering threats in areas defined as key such as autonomous systems, sensors, cyber and space.[9] However, as the 2014 UNIDIR report finds, humans can often have a poor record in foreseeing the full range of benefits and risks of new technologies.[10] Likewise, at this juncture, there still seems to be some uncertainty as to how such autonomous technologies will mature, and it is difficult to fully gauge the extent of the benefits and risks associated with their use.

In fact, it is unclear whether these advanced systems can in fact be developed. Several analysts even argue that assessments of capabilities and limitations of these systems are speculative to date, especially since there are no such weapons in the military environment yet because of operational and technological limitations.[11] They therefore consider that this debate is pointless since despite the rapid development of technology, fully autonomous weapons systems are still far off.[12]

Nevertheless, although the pace of improvement seems to be uncertain and there is significant disagreement on the state of the development of component technologies that are not equally advanced, in particular AI and machine learning, it seems that this is still an area of extremely high investment by both private and

---

4 Autonomy in this context is understood to be independence of control. Different factors allow for varying degrees of autonomy.

5 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 1.

6 Heyns, "Lethal Autonomous Robotics", UNIDIR Conference.

7 DCDC, *Strategic Trends Programme: Global Strategic Trends – Out to 2040*, UK Ministry of Defence, 4th ed., January 2010. See also: Caitríona Heinl, "Artificial (Intelligent) Agents and Active Cyber Defence: Policy Implications", 2014 6th International Conference on Cyber Conflict, NATO CCD COE.

8 Liran Antebi, "Who Will Stop The Robots"?, *Military and Strategic Affairs*, Volume 5 No.2, September 2013, 63. See also: UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 6.

9 HM Government, *Securing Britain in an Age of Uncertainty: The Strategic Defence and Security Review*, October 2010, 28.

10 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 8.

11 Gabi Siboni & Yoni Eshpar, "Dilemmas in the Use of Autonomous Weapons", *Strategic Assessment*, Volume 16 No.4, January 2014, 77.

12 Siboni & Eshpar, "Use of Autonomous Weapons", 76 & 82.

military sectors and it is not too early to discuss potential consequences.[13] In an effort to avoid such debate over dates, a recent Centre for a New American Security (CNAS) report calls this the 20YY regime since it argues that it will take some time for autonomous systems to become central to combat.[14] In this report, Robert Work, now U.S. Deputy Secretary of Defense, similarly posits that this is not the realm of science fiction, that a shift is coming and a slow recognition of these powerful trends will put tomorrow's [U.S.] military at unnecessary risk.[15]

Strategic futures reports forecast that growth in the role of unmanned, autonomous and intelligent systems is expected. A U.K. analysis of the future strategic context for defence, for instance, states that advances in robotics, sensors, energy efficiency, nanotechnology, and cognitive science coupled with powerful computing, will combine to produce rapid improvements in the capabilities of combat systems.[16] This analysis concludes that developments may be revolutionary where disciplines interact such as the combination of cognitive science and information communications technologies (ICT) to produce advanced decision-support tools. Therefore, quantum computing, simulation, AI, virtual databases, cognitive/behavioural science, and the reverse-engineering or mapping of the human brain, are relevant to advances in this space. Likewise, the revolutionary potential of future unmanned systems is tied directly to several interrelated rapid developments in the technology sector, especially trends in ICT, that will make unmanned systems increasingly capable, autonomous, and cost-effective, and these include computing power, cyber technologies, protected communications, big data, AI, autonomy, miniaturisation, commercial robotics, electric weapons, human performance modification, and additive manufacturing.[17]

Autonomous Technologies and Lethality

While UNIDIR's March 2014 report notes that several states have expressed interest in moving towards greater autonomy, perhaps as far as fully autonomous weapons, several delegations at a UNIDIR informal meeting of experts on lethal autonomous weapons systems in May 2014 indicated that there are no plans to develop such systems.[18] Some experts at this meeting further argued that there is little interest in deploying these systems to replace humans in an operational context. Several difficulties concerning autonomy arise with lethality and the application of increasing autonomy to functions such as target selection, the decision to use force and weapons release.[19]

Heyns describes this as a unique situation in that weapons are not just upgraded, but the weapon becomes the warrior, so a key issue of concern focuses on who is in fact making the decision to use force.[20] These discussions on the legal and ethical implications of such autonomous weapon systems first gained major public impetus following a Human Rights Watch (HRW)/Harvard Law School Human Rights Clinic (HLSHRC) position paper on the sale and use of autonomous weapon systems in November 2012 and a U.S. Department of Defense (DoD) Directive on autonomy in weapon systems.[21]

---

13 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 10.

14 Robert Work & Shawn Brimley, "20YY: Preparing for War in the Robotic Age", CNAS, January 2014, 6.

15 Work & Brimley, "War in the Robotic Age", 6.

16 DCDC, *Global Strategic Trends – 2040*. See also: Heinl, "Artificial (Intelligent) Agents and Active Cyber Defence".

17 Work & Brimley, "War in the Robotic Age", 8.

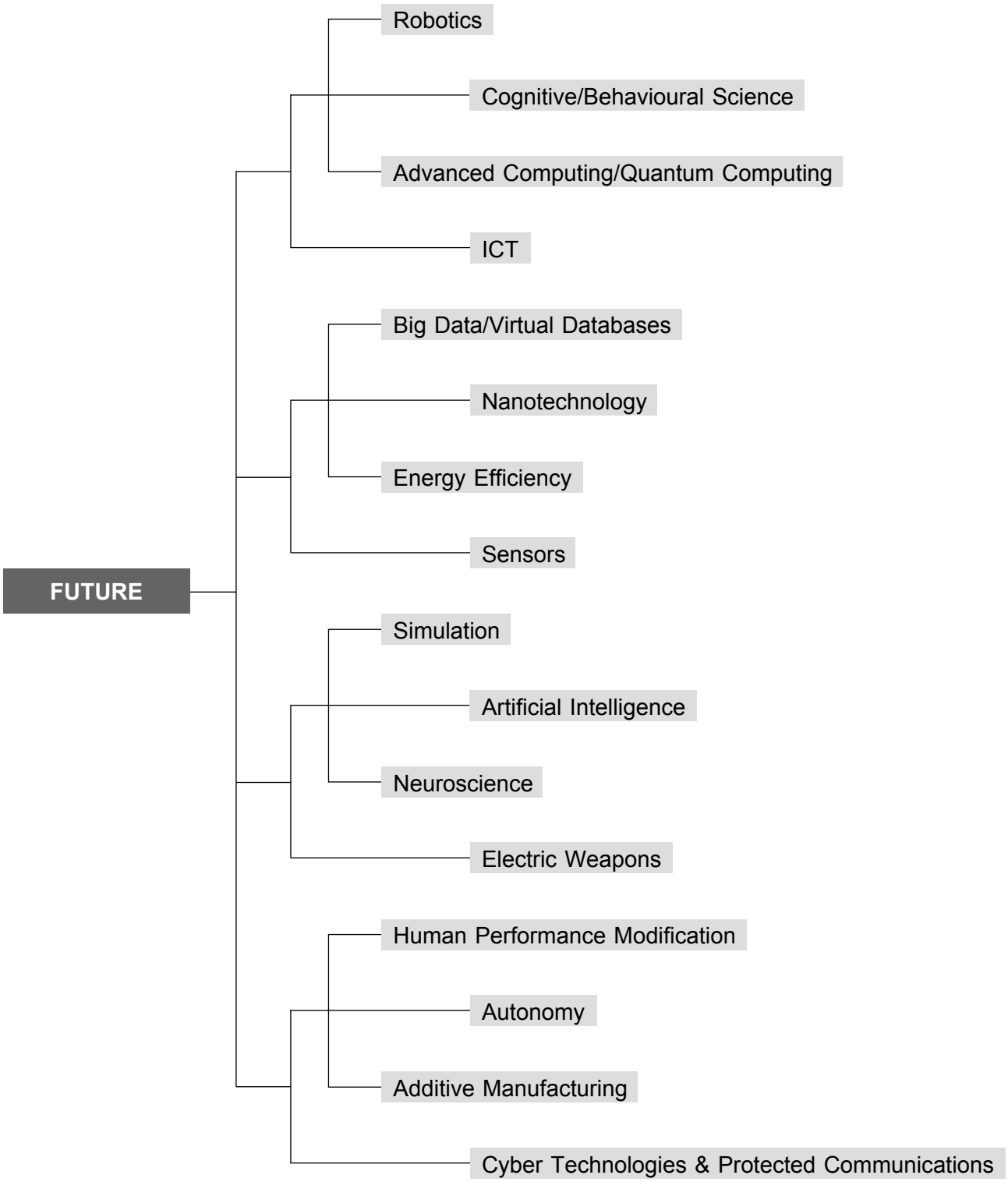
18 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 3.
See also: Chairperson of the Meeting of Experts, *Report of the 2014 informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, UNIDIR, 16 May 2014, 3.

19 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 3.

20 Heyns, "Lethal Autonomous Robotics", UNIDIR Conference.

21 Siboni & Eshpar, "Use of Autonomous Weapons", 80.

Scientific Areas of Strategic Importance:



The HRW/HLSHRC position paper called for an immediate stop to increasing autonomy by way of an international treaty to ban the development, sale and use of autonomous weapon systems. And, the U.S. DoD Directive on autonomy in weapon systems was issued in order to establish DoD policy and assign responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.[22] This Directive also suggests that guidance, established in line with the Directive, will be reviewed as necessary given the continual advancement of new technologies and changing warfighter needs. Subsequently, in examining the issue of autonomy in a report for the Human Rights Council in May 2013, Heyns recommended a freeze on efforts to develop autonomous weapons until an agreed international framework on their future is formulated.[23] He further suggested that a high-level panel of individuals from different backgrounds would be an important avenue for suggestions on a way forward.

Since then, 44 states have expressed apprehension over the challenges created by fully autonomous lethal weapons.[24] UNIDIR also launched a multi-disciplinary project to advance the multilateral discussion by refining areas of concern and identifying relevant research.[25] In November 2013, states party to the Convention on Conventional Weapons (CCW) adopted a decision to discuss for the first time in May 2014 how these challenges could be addressed, including assurance of meaningful human control over targeting decisions and the use of violent force. From Asia, China, India, Japan, Lao PDR, Pakistan, the Philippines, and the Republic of Korea, each party to the CCW, participated in this May 2014 informal meeting of experts to address questions related to emerging technologies in the area of lethal autonomous weapons systems in the context of the CCW's objectives. Indonesia, Malaysia, Myanmar, Singapore, and Thailand, although not party to the Convention, participated as observers.

Defining Autonomy

There seems to be still some divergence in the understanding of the nature of autonomous technologies. This section therefore examines several definitions for autonomy.

The informal meeting of experts in May 2014 concluded that the degree of autonomy can be defined by the level of human control on the system, or depend on the environment in which the system is supposed to operate, its functions, and the complexity of the tasks envisioned.[26] According to the UNIDIR 2014 report, autonomy can range from objects that are controlled by human operators at a distance to automatic/automated systems to fully autonomous systems.[27] Alternatively, autonomy could be divided into categories such as platforms controlled by human operators; platforms authorised by human operators; platforms supervised by human operators; and full autonomy.[28] Appendices 1 and 2 outline variables for autonomy in more detail. Surveillance devices, targeting devices, land-based vehicles, aerial vehicles, and robots are

---

22 United States Department of Defense, *Directive Number 3000.09*, 21 November 2012.

23 Siboni & Eshpar, "Use of Autonomous Weapons", 81.

24 Campaign to Stop Killer Robots, http://www.stopkillerrobots.org/2013/11/ccwmandate/.
Algeria, Argentina, Australia, Austria, Belarus, Belgium, Brazil, Canada, China, Costa Rica, Croatia, Cuba, Ecuador, Egypt, France, Germany, Ghana, Greece, Holy See, India, Indonesia, Iran, Ireland, Israel, Italy, Japan, Lithuania, Madagascar, Mexico, Morocco, Netherlands, New Zealand, Pakistan, Russia, Sierra Leone, Spain, South Africa, South Korea, Sweden, Switzerland, Turkey, Ukraine, United Kingdom, and United States.

25 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 2.

26 Chairperson, *2014 informal Meeting of Experts on LAWS*, 3.

27 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 2.

28 Antebi, "Who Will Stop The Robots?", 64-65.

examples of remote-controlled and automatic/ automated devices.[29] Whereas strategic futures reports suggest that systems could range from small sensors and personalised robots replicating human behaviour and appearance to a cooperative plethora of intelligent networks or swarms of environmental-based platforms with the power to act without human authorisation and direction with a range of autonomy from fully autonomous to significantly automated and self-coordinating while still under high-level human command.[30]

The DoD Directive defines an autonomous weapon system as "a weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation."[31] A human-supervised autonomous weapon system is defined as "an autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur". Whereas a semi-autonomous weapon system is defined as "a weapon system that, once activated, is intended to only engage individual targets or specific groups that have been selected by a human operator." Appendix 3 provides an outline of the definitions for these systems and their intended use according to the DoD Directive.

Heyns finds that while there is disagreement on the merits, the definition accepted by the U.S. DoD and HRW for lethal autonomous robotics includes weapons systems that once activated by a human, can engage and target individuals without further human intervention.[32] The findings of the informal meeting of experts in May 2014 also concluded that key elements for lethal autonomous weapons systems include the capacity to select and engage a target without human intervention.[33]

Nevertheless, delegations emphasised that discussions are at a very early stage and there is still a need to assess the current status as well as future trends in robotics. It is still unclear how this area will develop, thus there seems to be some unwillingness on the part of several states to make any commitments. The majority felt that the meeting assisted in understanding the characteristics of lethal autonomous weapons systems, but it is still premature to determine where discussions will lead. As a result, while the issue of a common definition for lethal autonomous weapons systems was raised and while some suggested that clarification would be required if more substantial work will be undertaken, most indicated that it is still too early to engage in such a negotiation.

Recommendations instead included exchange of information, development of best practices, a moratorium on research, and a possible ban. While the exchange of information and development of best practices are welcomed suggestions both in terms of informing the debate and developing confidence building measures, a moratorium or ban on research or further development for such technologies could prove extremely difficult to enforce. Nonetheless, although the meeting of experts resulted in informing a much-needed deeper common understanding of the issues, these

---

29 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 3.

30 DCDC, *Global Strategic Trends – 2040*.

31 Whereas a semi-autonomous weapon system is defined as "a weapon system that, once activated, is intended to only engage individual targets or specific groups that have been selected by a human operator."

32 Heyns, "Lethal Autonomous Robotics", UNIDIR Conference.

33 Chairperson, *2014 informal Meeting of Experts on LAWS*, 3.
Discussions at this meeting focused on technical matters, ethics and sociological issues, international humanitarian law, other areas of international law, and operational and military aspects.

recommendations are solely suggestions that are non-binding and many questions still remain unanswered in this area.

For instance, several military experts from the U.S. have begun to identify a clear need to account for this emerging set of new, potentially disruptive technologies, which may create sharp discontinuities in the conduct of warfare.[34] Similarly, several country delegations and experts at the informal meeting on lethal autonomous weapons systems in May 2014 described such systems as "a real game changer" in terms of military affairs.

Brain and Brawn: Military implications

"But the shift to something resembling guided munitions parity is only a predicate challenge to a potentially deeper revolution afoot – a move to an entirely new war-fighting regime in which unmanned and autonomous systems play central roles for the United States, its allies and partners, and its adversaries. U.S. defense leaders should begin to prepare now for this not so distant future – for war in the Robotic Age."[35]

The character of state-on-state conflict is changing and asymmetric tactics like economic, cyber and proxy actions, instead of direct military confrontation, are likely to play bigger roles as both state and non-state actors will try to seek an edge over those they cannot match with conventional military capability.[36] U.K. strategic defence and security reports note that the battlespace increasingly involves unmanned and cyber operations, and there are a number of capabilities such as weapons of mass destruction, emerging technologies with potential military application, and the systems used to deploy them, which could dramatically increase the risks of hostile acts should they reach the wrong hands.[37] That said, the importance of new weapons technologies can sometimes be overstated and the outcome of conflicts is not necessarily decided by technology but by the basic dynamics of the conflict itself, strategy, and political interests, and these factors should therefore also be taken into account.[38]

Given current geopolitical uncertainties in the Asia Pacific and tensions over territorial claims, such new technologies, if developed or acquired by state actors or malicious non-state actors, might only add to the complexity of these tensions. In light of the speed at which these technologies might be developed, it is also unlikely that policies or strategies will be in place in a timely or effective manner. Moreover, the mechanisms to prevent misunderstandings that might arise because of such emerging technologies are not yet in place. If states in the region consider that it might be less expensive to either develop or acquire these technologies relative to conventional weapons, it is likely that their increasing defence investments would also include such capabilities where possible. This is especially the case where countries might attempt to project influence that would otherwise be limited using conventional instruments.

Singapore, South Korea, Japan, and China lead in terms of advanced technologies in this region. The U.K. Ministry of Defence future reports also outline, for instance, that the scale and pace of the innovative and industrial capacity of countries like India and China will outpace many Western nations in a matter of years with China likely to attain and sustain global leadership in a number of technical areas including computer science.[39]

---

34 Work & Brimley, "War in the Robotic Age", 7.

35 Ibid, 5.

36 HM Government, *Strategic Defence and Security Review* 2010, 16.

37 Ibid, 55.

38 Chairperson, *2014 informal Meeting of Experts on LAWS, 5*. See also: UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 7.

39 DCDC, *Global Strategic Trends – 2040*.

Although it is not yet clear how, or whether, this would indeed affect the regional balance of powers, Robert Work, now U.S. Deputy Secretary of Defense, and Shawn Brimley, former Director for Strategic Planning on the White House National Security Staff, have previously recommended that new norms of behaviour will need to be developed as leaders adapt to the unique attributes and challenges of autonomous systems in crisis situations.[40] Furthermore, non-state actors could also further complicate matters since terrorist groups, organised crime gangs or proxy actors might even obtain or alter commercially available technologies. This aligns with projections that the character of war is likely to continue to be shaped not only by a system of rival states, but by forces outside the state-centric systems.[41]

Defence reports assert that militaries will seek superior technology, including the most advanced civilian technology that can be adapted, to gain operational advantage (although this will differ from those available against a non-state actor).[42] Likewise, given the military build up in the Asia Pacific, it is likely that states will seek such superior technology to gain an edge over their potential adversaries. Military experts argue that remotely piloted air and ground vehicles will soon be replaced by increasingly autonomous systems in all physical operating domains (air, sea, undersea, land and space) and across the full range of military operations.[43] Currently, it seems that military interest lies in a limited range of missions such as force protection, demining, surveillance of dangerous environments, and defensive use to protect borders or military installations, via sensors or systems capable of attack.[44] Robert Work and Shawn Brimley further argue in their CNAS report that the U.S. will be driven to these systems out of operational necessity, as well as the costs of personnel and the development of traditional crewed combat platforms that are increasing at an unsustainable pace.[45]

In terms of potential military advantages so far identified, these seem to include reduced risks to soldiers' lives, the ability to undertake tasks that humans cannot perform due to physical limitations, greater force projection, freeing humans from repetitive tasks, greater endurance, better precision, faster information processing, more direct targeting, and a belief by some that these systems will one day respect international humanitarian law and human rights law better than humans, especially if they also have greater powers of distinction.[46] Moreover, these tools may be deployed very quickly and they can react in a short space of time, especially if humans are not in the decision loop when speed is of the essence.[47] Autonomous technologies could also possibly assist rescue missions, protection of armed forces and civilians, logistics, transportation, intelligence, law enforcement authorities, and peacekeeping missions.

Another major argument for the use of advanced automation and increasingly autonomous technologies is the possible alleviation of constraints on labour and financial resources. Although the development and procurement of some of these technologies might be expensive, significant economic advantages could include lower army maintenance

---

40 Work & Brimley, "War in the Robotic Age", 31.

41 Pascal Vennesson, "Dimensions of War and Strategy", 15th Asia Pacific Programme for Senior Military Officers – The Future of War, RSIS Singapore, 5 August 2013.

42 UK Ministry of Defence, *National Security Through Technology: Technology, Equipment, and Support for UK Defence and Security*, February 2012, 26.

43 Work & Brimley, "War in the Robotic Age", 6.

44 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 6.

45 Work & Brimley, "War in the Robotic Age", 6.

46 Antebi, "Who Will Stop The Robots?", 61. See also: UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 6.

47 Heyns, "Lethal Autonomous Robotics", UNIDIR Conference.

costs or a smaller number of operators' salaries.[48] Furthermore, most states are already challenged in identifying, training and retaining large numbers of skilled individuals in ICT and cybersecurity for example. Not only is burnout in the current cyber workforce already beginning to show according to U.S. defence reports, expected demographic trends could even work against several countries.[49]

Expected demographic shifts are another significant factor that should be taken into account. Developed countries face rapidly aging populations and falling birth rates, while emerging economies can expect more working-age adults due to rising birth rates, and these changes are expected to have a dramatic effect on resource needs and long-term economic sustainability.[50] Some countries like Japan already have disproportionately large populations of older people and this trend will become more prominent.[51] Therefore, while some countries may consider autonomous weapons systems a response to manpower crises, others might not face the same shortage, and this should be considered when analysing how military interest in autonomy could evolve.[52]

In terms of disadvantages, there are risks such as vulnerability to cyber attacks, a lack of predictability, and difficulties in adaptation to complex environments.[53] These systems could also have consequences for arms control. Military strategists from different countries further question the necessity or desirability of delegating responsibility for a decision on launching an attack to autonomous systems since the unacceptable political costs for incorrect action might be too risky to delegate to machines.[54] Robert Work and Shawn Brimley write for instance that technology does not make war more clinical; it makes it more deadly, and precision does not make the battlefield more sterile but rather makes it increasingly lethal.[55]

The asymmetric use of lethal autonomous robots on a battlefield could even mean that rather than fighting robots, people might instead attack civilian populations.[56] Because such technology might allow leaders to engage in conflicts without risking soldiers' lives, this could mean that military options might be chosen over a policy of dialogue and avoidance of conflict.[57] Managing stability in periods of tension might therefore become more difficult, and such systems could reshape how the U.S. military bases its forces around the world or how decisions are made by policymakers about the use of force.[58] For example, while U.K. strategic defence and security government reports stress that current thinking suggests that it costs far more when conflict is not prevented and government intervenes militarily, such technologies might alter this kind of strategic thinking.[59]

---

48 Antebi, "Who Will Stop The Robots", 67.

49 Under Secretary of Defense, *Resilient Military Systems*. See also: William Lynn III, former United States Under Secretary of Defense, "2010 Cyberspace Symposium – DoD Perspective", 26 May 2010.

50 Microsoft, "Cyberspace 2025 – Today's Decisions, Tomorrow's Terrain: Navigating the Future of Cybersecurity Policy", June 2014, 10.

51 Ibid, 11.

52 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 7.

53 Chairperson, 2014 informal Meeting of Experts on LAWS, 5.

54 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 6.

55 Work & Brimley, "War in the Robotic Age", 9.

56 Heyns, "Lethal Autonomous Robotics", UNIDIR Conference.

57 Siboni & Eshpar, "Use of Autonomous Weapons", 81.

58 Work & Brimley, "War in the Robotic Age", 9.

59 HM Government, *Strategic Defence and Security Review 2010*, 3.

At least publicly, it does not seem that the implications of such scenarios have been examined at great length yet, nor have strategic countermeasures been created to deal with the possibilities of such negative outcomes. In Israel, analysts argue that the nation has a clear interest in promoting local and international mechanisms that will give legitimacy to the use of autonomous capabilities in weapon systems within the framework of the ethical restrictions to which it is committed.[60] They conclude that the integration of these capabilities into weapon systems can potentially bring great military benefit while meeting accepted legal standards and sometimes even meeting higher moral standards.[61] But harnessing these operational and ethical benefits will depend on developing legal and political tools that will effectively curb dangerous technological developments and prevent immoral use.[62] While in the U.S., defence experts argue that a warfare regime based on unmanned and autonomous systems has the potential to change basic concepts of defence strategy including deterrence, reassurance, dissuasion, and compellence, as well as military concepts such as the relationship between offensive and defensive military strategies or the interplay of range, mass and speed.[63] If this is in fact the case, there is then an urgent need to begin developing alternative strategies that will take these unique factors into account.

In some respects, human-machine collaborations might become the preferred solution. AI, for example, may never be as powerful as "intelligence amplification", which is when human cognition is augmented by close interaction with computers.[64] By way of example, when a machine and human chess player were paired in collaboration, tests found that human-machine teams, even when they did not include the best grandmasters or most powerful computers, consistently beat teams composed solely of human grandmasters or computers.[65] Likewise, the UNIDIR report of 2014 also recommends further exploration of the strengths of both man and machine, and how the benefits of certain uses of autonomy could be harnessed without sacrificing humanity.[66] IBM authors on smart machines write that the goal is not to replicate the human brain or replace human thinking with machine thinking. Rather, in the era of cognitive systems, humans and machines could collaborate to produce better results – each bringing their own superior skills to the partnership. Machines will be more rational and analytic, possessing encyclopaedic memories and tremendous computational abilities, whereas individuals are expected to provide judgement, intuition, empathy, a moral compass and human creativity. The CNAS report similarly concludes that "[t]he 'winners' will likely be those who best leverage the unique advantages of both machine and human intelligences".[67]

Cyber-Related Implications

So far discussions seem to have focused to a lesser extent on the cyber-related implications of these technologies. The degree of vulnerability due to the underlying systems needs further examination since, for example, tools could be taken over or information intercepted by cyber means.[68] Similarly, Robert Work's

---

60 Siboni & Eshpar, "Use of Autonomous Weapons", 77.

61 Ibid, 84.

62 Ibid

63 Work & Brimley, "War in the Robotic Age", 6.

64 Walter Isaacson, "Brain gain?", Book Review of Smarter Than You Think by Clive Thompson, *International New York Times*, 2-3 November 2013. See also: Heinl, "Artificial (Intelligent) Agents".

65 Isaacson, "Brain gain?". See also: Heinl, "Artificial (Intelligent) Agents".

66 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 7.

67 Work & Brimley, "War in the Robotic Age", 25.

68 Antebi, "Who Will Stop The Robots", 70.

CNAS report foresees that cyber operations are a rapidly advancing dimension that will intersect heavily with warfare in the robotics age, and cyber is likely to be the new high ground in future warfare, particularly since an actor who dominates in cyber conflict can potentially shut down or usurp control over physical platforms which is especially true for unmanned systems.[69]

UNIDIR's report also identifies a lack of critical analysis on whether increasingly autonomous weapons systems will drive development of other weapons, countermeasures or methods including cyber conflict.[70] For instance, it is unclear how such technologies might impact the effectiveness of the nuclear deterrent as the ultimate means to deter the most extreme threats. It is probable that actors would seek out vulnerabilities in such systems and it is therefore recommended that a more thorough examination be undertaken on how such countermeasures or cyber tools might consequently develop.

Cyber systems do not seem to be mentioned extensively within the current deliberations on these technologies. For instance, it is not very clear whether the U.S. DoD Directive applies to fully autonomous or semi-autonomous systems for cyberspace operations. Intelligent software is used increasingly in cyber operations and several analysts are arguing that defence systems be even more adaptive and evolve dynamically with network conditions changes by implementing dynamic behaviour, autonomy, and adaptation such as autonomic computing or multi-agent systems.[71] Such autonomous intelligent agents can be purely software operating in cyberspace (computational agents) or integrated into a physical system (robotic agents) where they underpin a robot's behaviour and capabilities.[72] Since intelligent agents can seemingly be used most efficiently in multi-agent formations, it is expected that this will be the main form of agent application in cyber operations.[73]

However, significant questions are already being raised by a number of scientists and policy analysts that require further concrete examination. For instance, intelligent agents in multi-agent formations could apparently negotiate between themselves and cooperatively behave in a complex way to achieve a commander's general goals, but strict control of each single agent's behaviour will be weaker and it could be impossible to verify the outcome of multi-agent behaviour for all situations. If agents have too much autonomy in decision-making, unwanted coalitions might occur since their communication would only be partially visible to human controllers and this might be very difficult to disable.[74] The more intelligent software becomes, the more difficult it might be to control. Furthermore, there are several related challenges including the complexity of agents' behaviour, misunderstanding situations, misinterpretation of commands, loss of contact, and formation of unwanted coalitions, unintentionally behaving in a harmful way or unexpected actions and unpredictable behaviour.[75] In addition, advanced intelligent systems could challenge the interaction between automated and human components, and the complexity of controlling multiple autonomous systems and interpreting information could become extremely difficult. U.K. MoD strategic

---

69 Work & Brimley, "War in the Robotic Age", 23.

70 UNIDIR, "Weaponization of Increasingly Autonomous Technologies", 8.

71 Igor Kotenko, "Agent-based modelling and simulation of network cyber-attacks and cooperative defence mechanisms", St. Petersburg Institute for Informatics and Automation, Russian Academy of Sciences, http://cdn.intechopen.com/pdfs/11547/InTech-Agent_based_modeling_and_simulation_of_network_infrastructure_cyber_attacks_and_cooperative_defense_mechanisms.pdf, 2010.

72 Guarino, "Autonomous Intelligent Agents".

73 Tyugu, "Command and Control of Cyber Weapons".

74 Tyugu, "Command and Control of Cyber Weapons". See also Guarino, "Autonomous Intelligent Agents".

75 Tyugu, "Command and Control of Cyber Weapons". See also: Heinl, Artificial (Intelligent) Agents.

futures forecasts therefore suggest that those unsuitable for these challenges may be replaced by intelligent machines or "upgraded" by technology augmentation.[76]

In conclusion, while there is still some uncertainty as to how maturing autonomous technologies, including potentially fully autonomous and lethal systems, will develop and impact national security, it is clear that several major policy questions are already evident. Key questions are identified throughout both parts of this report, which should be considered given the increasing interest in these technologies from both military circles and industry.

This first part of the report finds that a clearer understanding of the nature of these technologies would assist this debate, especially since it is likely that states will pursue technological superiority via increasingly autonomous technologies for both economic and military reasons. Deeper analysis is required on the possible military advantages and disadvantages that might ensue, including the role of the human vis-à-vis the machine.

The second part of the report finds that currently, there are also major challenges in controlling and regulating this space as well as highly significant legal ambiguities and ethical question marks. It argues that the relationship between the public sector and industry should also be better managed to ensure that while innovation and economic growth are not restrained, this area will be developed responsibly. Lastly, while policy guarantees that the operation of systems will always be under human control, it does not seem certain from a technical standpoint that the human might always be in a position to control such systems.

---

[76] DCDC, *Global Strategic Trends – 2040*.

Appendix 1

Categories of Autonomy:

Adapted from the article, "Who Will Stop the Robots?", Liran Antebi[77]

Platforms controlled by human operators	The human operator makes all the decisions. The system has no independent control over its environment e.g. toy car operated by remote control.
Platforms authorised by human operators	The platform performs actions independently when it is authorised to execute them by a human operator e.g. robotic vacuum cleaners that when turned on receive authorization to clean without outside intervention.
Platforms supervised by human operators	The system can carry out a wide range of actions independently when it receives the approval or instructions from a human operator. Both the human operator and the system can begin an action based on information received from sensors but the system can do so only within the range of tasks that it is planned to carry out.
Full autonomy	The system receives targets from human operators and translates them into tasks that will be performed without any human intervention including the stage of planning and choosing the means of implementation. The human operator can still intervene and influence events when necessary.

[77] Liran Antebi, "Who Will Stop The Robots"?, Military and Strategic Affairs, Volume 5 No.2, September 2013, p.64/65.

Appendix 2

Variables Relevant to Legality and Acceptability:

Acceptability of a system could be informed by the interactions of these variables since they have direct impact on considerations of legality and acceptability.

Adapted from UNIDIR Resources, "Framing Discussions on the Weaponization of Increasingly Autonomous Technologies", March 2014.

Goal-satisfying actions	The ability to create and follow plans of action aimed at satisfying goals. Are goals generated by the system itself or determined by an external source ("orders")? Are plans generated by the system and then vetted through external confirmation (seeking approval) or simply implemented?
Predictability	Predictability of actions the system may take. The simpler the environment, the less the need for a variety of actions and the more predictable a system will become. Likewise, the less variability in the type of actions a system can take, the more predicable it will be, even in a complex environment. The tighter the control that is applied to a system (for example, unwavering focus on a specific goal) the less variable and more predictable a system will be.
Communication	How precise does communication with the system need to be (i.e. does decreasing precision in communication mean the system has to increasingly "interpret" meaning)? How frequent is its communication?
Depth of Reasoning	The more limited the reasoning a system is capable of, the less in the way of autonomy it will have and the more predicable it will appear. A simple environment requires less depth of reasoning than a more complex one.
Precision of Sensors and Capacity for Synthesis	The raw sensory capabilities of a system will determine its ability to discriminate things in its environment. The ability to combine different types of sensors and synthesize a view of the environment provides a more refined basis for situational awareness.
Bounds on Location or Operating Environment	Control of the physical location and the complexity of the environment in which a system may function will increase control over a system.
Functions	The nature of the actions available to a system (for example navigation, targeting, or weapons release).

Appendix 3

U.S. Department of Defence Directive: Framework

System	*Definition*	*Intended Use*
Semi-autonomous weapons systems (including manned or unmanned platforms, munitions, or sub-munitions that function as semi-autonomous weapon systems or as subcomponents of semi-autonomous weapon systems)	A weapon system that, once activated, is intended to only engage individual targets or specific groups that have been selected by a human operator.	May be used to apply lethal or non-lethal, kinetic or non-kinetic force. Those that are onboard or integrated with unmanned platforms must be designed such that, in the event of degraded or lost communications, the system does not autonomously select and engage individual targets or specific target groups that have not been previously selected by an authorised human operator.
Human-supervised autonomous weapon systems	An autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur.	May be used to select and engage targets, with the exception of selecting humans as targets, for local defence to intercept attempted time-critical or saturation attacks for: a) static defence of manned installations; b) onboard defence of manned platforms.
Autonomous weapon systems	A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.[78]	May be used to apply non-lethal, non-kinetic force, such as some forms of electronic attack, against materiel targets.[79]

[78] Whereas a semi-autonomous weapon system is defined as "a weapon system that, once activated, is intended to only engage individual targets or specific groups that have been selected by a human operator."

[79] In accordance with DoD Directive 3000.3.

About the Author

Caitríona H. Heinl is a Research Fellow responsible for research on cybersecurity matters under the Homeland Defence Programme at the Centre of Excellence for National Security (CENS) within the S. Rajaratnam School of International Studies (RSIS). CENS is a research unit which works closely with the National Security Coordination Secretariat (NSCS) within the Prime Minister's Office, Singapore.

About the Centre of Excellence for National Security

The **Centre of Excellence for National Security (CENS)** is a research unit of the S. Rajaratnam School of International Studies (RSIS) at the Nanyang Technological University, Singapore.

Established on 1 April 2006, CENS raison d'être is to raise the intellectual capital invested in strategising national security. To do so, CENS is devoted to rigorous policy-relevant analysis across a range of national security issues.

CENS is multinational in composition, comprising both Singaporeans and foreign analysts who are specialists in various aspects of national and homeland security affairs. Besides fulltime analysts, CENS further boosts its research capacity and keeps abreast of cutting edge global trends in national security research by maintaining and encouraging a steady stream of Visiting Fellows.

About the S. Rajaratnam School of International Studies

The **S. Rajaratnam School of International Studies (RSIS)** is a professional graduate school of international affairs at the Nanyang Technological University, Singapore. RSIS' mission is to develop a community of scholars and policy analysts at the forefront of security studies and international affairs. Its core functions are research, graduate education and networking. It produces cutting-edge research on Asia Pacific Security, Multilateralism and Regionalism, Conflict Studies, Non-Traditional Security, International Political Economy, and Country and Region Studies. RSIS' activities are aimed at assisting policymakers to develop comprehensive approaches to strategic thinking on issues related to security and stability in the Asia Pacific.

For more information about RSIS, please visit www.rsis.edu.sg.

RSiS
Nanyang Technological University
S. RAJARATNAM SCHOOL OF INTERNATIONAL STUDIES

Nanyang Technological University
Block S4, Level B4, 50 Nanyang Avenue, Singapore 639798
Tel: +65 6790 6982 | Fax: +65 6794 0617 | www.rsis.edu.sg