

BEYOND ASIMOV'S LAWS OF ROBOTICS: SLEEPWALKING TOWARD THE FUTURE?

On 16 May 2014, the first multilateral discussion of lethal autonomous systems was convened at the United Nations Office in Geneva on the margins of the Expert Meeting for the Convention on Certain Conventional Weapons. The subject of discussion, in more colloquial terms, was the so-called 'killer robot' – that is, fully autonomous lethal weapons systems that can select and engage targets without human intervention. During this milestone meeting, discussions touched on the various technical, legal, operational and ethical implications of robotic weapons and, of course, a number of concerns were voiced about a future with weapons void of human control, judgment or compassion.



A particularly descriptive statement came from the Permanent Representative of Brazil to the Conference on Disarmament, His Excellency Pedro Motta, who drew an analogy between killer robots and the old Jewish tale of the Golem (1). According to the tale, the Golem was built from clay as a tool to defend the Jewish community of Prague from those who would wish them harm or wrongdoing. At first it served the community well. One day, however, the Golem went out of control and murderously rampaged through the streets until its creator, the Rabbi, could once again regain control and return the Golem to inanimate clay. The point to be taken from this old tale, His Excellency observed, is that society should be careful when relying on technology as a solution for its challenges.

While it is commendable that the issue of robotics is finally beginning to appear on the international agenda as a matter of safety and security, the present day reality with regard to the governance of robotics leaves a lot to be desired. This brief discourse seeks to demonstrate this concern and to shed light on the need to figuratively 'wake' from our sleep with regard to robotics. In doing so, we will turn to two existing international instruments that have succeeded in addressing the sensitive and divisive issue of non-proliferation of chemical and nuclear weapons, with a view towards learning some useful lessons to shape a future system of governance of autonomous robotics.

Robots of the present - Science fiction no more

Although such fully autonomous weapons systems, or killer robots, are not currently in use, research and

development in all aspects of autonomous robotics is rapidly expanding. Having left the pages of science fictions books, robots with varying degrees of autonomy have begun to penetrate our daily lives. iRobot's Roomba (2) vacuums your floors, Robomow's robot mows your lawn (3) and JIBO (4), the 'World's First Family Robot', entertains the kids, reads you your mail and orders you pizza online. If you live in Russia, a robotic drone can even deliver the pizza to your door when it is ready (5). Looking at the fast paced expansion of robotics, it is really not beyond belief that fully autonomous robots could be in wide spread use, as weapons of war, within 20 to 30 years. A 2003 report from the U.S. Joint Forces Command already envisages that "between 2015 and 2025 the joint forces could be largely robotic at the tactical level" (6). Of course, there is no widespread agreement on the above prediction and some consider this to be a greatly exaggerated future (7). The fact of the matter though is that, while we should try to remain as grounded in reality as possible and avoid whimsical speculation steeped in science fiction, the danger of being wrong about these predictions is too great a risk to take.

Noel Sharkey, Professor of Artificial Intelligence and Robotics at the University of Sheffield, expressed his belief that truly autonomous robotics are not realistic in the immediately near future. However, he expressed his concern that nonetheless "we are in danger of sleepwalking into a world where robots have become indispensable to service our needs, to care for us, to watch our every activity, to police us and to fight our wars for us." (8) For these reasons alone, it would be prudent to start paying due attention to developments in the robotics world and to start considering how to start structuring a system of governance of robotics for the future.

In the works of renowned science fiction author, Isaac Asimov, the system of governance was very straightforward as the potential dangers of robotics were famously mitigated by three laws of robotics that were programmed into Asimov's robots:

- Law One: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- Law Two: A robot must obey orders given to it by human beings except where such orders would conflict with the First Law.
- Law Three: A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law (9).

Although nice on paper, the laws have not materialised in reality because modern day roboticists have not been able to translate them from a fictional plot tool into the language of robotics. Even if these laws were somehow implementable in real life, there is the further problem that they are even too focused on physical harm to be able to really deal with some of the other modern concerns for robotics, such as issues of privacy stemming from data gathering or tracking by domestic robots. Nonetheless, Asimov's Laws, limited as they may be, are all we have to show for the moment and little or no progress has been made towards a system of governance by policy makers in this particularly technical field. In this regard, society is on the edge of a 'wild west' of robotics, where the only laws of robotics in existence are those found in the pages of science fiction short stories.

Looking for something more concrete to guide us towards and through the coming era of robotics, many involved in the killer robot discussions under the Convention on Certain Conventional Weapons support an approach based on the belief that the best way to prevent the Golem's rampage may in fact be by not creating it in the first place. To achieve this, some have proposed the adoption of an international agreement banning the development, production and use of killer robots (10), while others have called for a moratorium on killer robots pending a greater understanding of their nature and relationship to international humanitarian law (1). Shedding light on the logic behind such a stance, one researcher at Princeton University jokingly noted that "we already don't understand Microsoft Windows, we're certainly not going to understand something as complex as a humanlike intelligence. Why should we create something like that and then arm it?" (12) While there may be some degree of merit in this logic, those against a ban or moratorium on lethal autonomous weapons systems believe that not enough thought has been given to the possible negative implications of a ban or moratorium on killer robots may have.

The reason for this is that the weapons side of robotics is but a part of the much larger and growing robotics industry. The military application of robotics is very much linked with the domestic and industrial application of robotics. Take for example, the aforementioned Roomba, which is designed by iRobot - the same company that produces PackBot, a versatile robot designed for use by military and law enforcement personnel for dangerous activities such as CBRN reconnaissance. iRobot is also partnered with Taser International, a leading manufacturer of stun guns (13). Similarly, Boston Dynamics and SCHAFT Inc. were recently acquired by Google as part of its 'moonshot' to create a new generation of robots. (14) In the not so recent past though, both Boston Dynamics and SCHAFT have also had close associations with the military application of robotics. Boston Dynamics' LS3, for example, was recently field tested by US Marines during the Rim of the Pacific Exercise — a multinational maritime exercise — and the SCHAFT robot succeeded in winning the DARPA Robotics Challenge 2013 Trial (15). DARPA is the Defense Advanced Research Projects Agency (DARPA), a technology research agency of the United States Department of Defense. Viewing domestic, industrial and military robotics thus interlinked, the response of the international community to the killer robot question, will impact on research and development of domestic and industrial robotics. How to proceed in this heated debate should therefore not be considered so straightforward. Instead, we should give due thought to the possibility that perhaps neither a ban on killer robots, under the Convention on Certain Conventional Weapons, nor the development of a new international instrument to this end, may not be the most appropriate solution to the challenge: both will have repercussions for the budding autonomous robotics industry and the robots that are becoming increasingly important for our daily lives.

Moving beyond Asimov - Lessons from the past

In attempting, therefore, to design a system of governance for robotics there are many examples of best practices in existing international instruments. These best practices would allow us to comprehensively take into account the complexities of the autonomous robotics world, while at the same time, providing enough structure to comfort those who express alarm over possible proliferation risks of killer robots. These instruments could be looked at as inspiration for equitably dealing with those with an interest in the future of robotics and in awakening us from our "sleepwalking".

Two possibilities that deal with sensitive and divisive subject matter, and are almost universally endorsed in doing so with 190 state parties each, are: the Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction (CWC) and the Treaty on Non-Proliferation of Nuclear Weapons (NPT). While both instrument have been extremely successful in terms of achieving an almost universal level of application by countries, few international instruments have achieved the recent level of success demonstrated by CWC in the destruction of chemical weapons stockpiles. (16) For this reason, while best practices can be drawn from both instruments, particular attention should be paid to the lessons learned from CWC (17).

The history of CWC demonstrates that reaching the objective of the complete elimination of the chemical category of weapons of mass destruction was not an easy task and that the lobby of the chemical industry — the major stakeholder of the international norm — was one not to be taken lightly. (18) Indeed, the opposition presented by the chemical industry resulted in a long and drawn out period of discussion and negotiation that started as early as the 1960's. Once CWC was eventually signed in 1993, it was still a further four years before CWC finally entered into force in 1997. Even then, the initial operative years were quite slow. The reasons behind this are well described by Mr. René van Sloten — Executive Director on Industrial Policy at the European Chemical Industry Council (CEFIC) — who underlined that the chemical industry manufactures products that "the whole world depends upon every day for health, safety, transportation, communication, agriculture, medicine - touching virtually every aspect of our lives. Chemical companies do not make chemical weapons, yet some very common chemicals can be misused as - or transformed into - chemical weapons." (19)

This is the so-called dual-use nature of certain chemicals. Chlorine, for instance, is a common example of a dual-use chemical; it can be used peacefully as a substance to purify water thereby making it safe for drinking or it can be used in its gaseous form as a chemical weapon. Evidently, the chemical industry had plenty grounds for concern with a Convention that seeks the elimination of chemical weapons. At the same time though, noting the importance of having the chemical industry onboard and supporting CWC, Mr. van Sloten

indicated that the “industry best practice in chemical site housekeeping and customer vetting can be used to provide a high degree of compliance and confidence in response to the non-proliferation aspects of the Convention.” (20)

In order to ensure, therefore, that the chemical industry was not alienated by overbearing terms and conditions, and to rely on the strengths of the industry to foster compliance, CWC was drafted with an approach that balanced the interests of the chemical industry and the need to address chemical weapons proliferation. This is largely done by accurately refining the scope of the Convention to ensure that ‘scheduled’ chemicals, as well as those chemicals that fall under the general purpose criterion, are only used for peaceful purposes and not for purposes prohibited by the Convention. (21) By doing so, the drafters of CWC ensured that the implementation of the Convention would not hinder the chemical industry in any manner - an act that would secure the full support of the industry allowing CWC to reach an almost universal application.

At present, the Convention enjoys enormous support marked by strong cooperation between the chemical industry and CWC through the Organisation for the Prohibition of Chemical Weapons (OPCW) - the CWC implementing body. His Excellency Minoru Shibuya, Ambassador of Japan and Chairman of the Conference of the States Parties to the Chemical Weapons Convention, described this cooperation during the 2009 Helsinki Chemicals Forum noting that “this is a unique and mutually beneficial partnership which is essential in promoting the full implementation of the Convention’s provisions. During the negotiations on the Chemical Weapons Convention, the chemical industry agreed to a ‘declaration and verification process by inspection’ to provide assurances that toxic chemicals and dual use chemicals are not used for weapons production. This was a major commitment and represents the backbone of the non-proliferation provisions in the Convention.” (22)

Indeed, the verification process is a further example of how CWC has managed to balance both sides to make a practical and effective solution. Adopting specific guidelines for verification that regulate inspection time and access to information and facilities, in a manner that respects the industry’s concern for confidentially handling sensitive commercial information, has gone a great distance to facilitate inspections by OPCW that the industry may ordinarily consider intrusive.

Reading into this comparison, a first step for determining how to proceed with killer robots — and by implication autonomous robotics in general — could be to shift the paradigm of the discussion away from viewing robotics as purely an issue of arms control and proliferation. Instead, robotics could be considered as a resource and discussions could be focused on the purposes assigned to the use of autonomous robotics as per CWC. In doing so, initiative could then be taken to build a collaborative partnership with the robotics industry. To facilitate this transition, the moniker of “killer robots” should be dropped, in place of the broader and more industry friendly term “autonomous robotics”. The logic behind this paradigm shift stems from the fact that autonomous robotics, like chemicals, are not inherently bad. Both are tools to which a purpose is applied. Both are dual-use.

Another best practice that could be taken from CWC is the role played by OPCW as a leading forum in the chemical field. Article IX of the Convention sets out the possibility for states party to the Convention to pursue consultation, cooperation and fact-finding missions through OPCW. This Article can be used by Member States for consultative purposes to ensure compliance and, in doing so, it adds transparency and credibility to the workings of the Convention. At the same time, Article XI acknowledges the need for economic and technical development and therefore explicitly states that the Convention should be implemented in a manner that does not hamper this or international cooperation to this effect, including “the international exchange of scientific and technical information and chemicals and equipment for the production, processing or use of chemicals for purposes not prohibited under the Convention.”

While these articles put OPCW in a pivotal position as the central hub for exchange of knowledge, best practices and lessons learned in the chemical realm, it must be admitted that they have not been exploited to the fullest extent by OPCW and the states party to CWC. That being said, the idea behind these articles could be applied to the future system for autonomous robotics and, if the immediately apparent benefits of such a

system for exchange are promoted properly, it could be an ideal mechanism for building a partnership with the autonomous robotics industry. As a field of scientific research, communication amongst roboticists is an essential aspect of the scientific method and, in this regard, the establishment of a forum for consultations, cooperation and clarifications on autonomous robotics would only serve to further support all those engaged and drive innovation. Accordingly, there are grounds to believe that such an approach would be openly greeted and embraced by the robotics industry, allowing it and any associated system of governance to thrive. As was the case with CWC though, steps would be required to ensure that the industry's commercially sensitive information is protected: a necessary precondition if this lesson learned is to work for autonomous robotics.

Evidently, the success of the CWC story has many lessons that could be of benefit to the ongoing killer robots debate. Its hallmarks of striking a balance between the needs of the industry and the needs of non-proliferation efforts, as well as fostering a continuous dialogue and collaborative relationship, have helped OPCW progress towards its mandate and even earn a Nobel Peace Prize. Sentiments expressed by OPCW Director-General, Mr. Ahmet Üzümcü, in 2012 again illustrate how this has been possible - "the implementation of the Convention is sufficiently nimble and robust, not only to keep pace with developments in the industry and in science and technology, but also to stay ahead of the curve." (23)

Looking next at NPT, the treaty first entered into force in 1970 — with an initial duration of 25 years — to prevent the spread of nuclear weapons and nuclear weapons technology. (24) In 1995 a review conference was convened which led to the treaty being extended indefinitely. There is little doubt that, since 1970, NPT has garnered widespread support making it the very corner stone of nuclear non-proliferation and has resulted in a declining rate of nuclear weapons proliferation since its adoption. This remains the case in spite of the usual critique of the unfairness or discrimination inherent in the fact that the treaty is founded on the concept of nuclear "haves" and "have nots" and that the five states permitted to possess nuclear weapons under the treaty still maintain considerable stockpiles of nuclear weapons to the present day. In this regard, there are certainly lessons that can be gleaned from how NPT and its associated organisation, the International Atomic Energy Agency (IAEA), have developed and maintained a position of almost universal support from 190 countries over the years.

One of the most important aspects of NPT is the third pillar of the Convention, which is embodied in Article IV and sets out the inalienable right of states party to the Convention to develop, research, produce and use of nuclear energy for peaceful purposes. Over the years, this right has facilitated more than thirty states to make nuclear energy a key component of the national power supply. At the same time, Article III of NPT envisages the development of a system of safeguards by IAEA to compliment this right by preventing the diversion of nuclear material from peaceful to weapon-based uses and expressly requiring non-nuclear weapon state parties to adhere to this system. IAEA uses this safeguard mechanism to verify declarations made by states regarding their nuclear activities and to detect, in advance, any potential misuse or violation of the Convention.

Similar to the objectives behind many provisions in CWC, Article III of NPT states that the safeguards shall be implemented in a manner that complies with Article IV and avoids "hampering the economic or technological development of the Parties or international cooperation in the field of peaceful nuclear activities, including the international exchange of nuclear material and equipment for the processing, use or production of nuclear material for peaceful purposes."

Looking at these fundamentals of the nuclear non-proliferation, it is evident that the treaty was drafted in a manner so as to carefully cater to the needs and interests of both sides: those who acknowledge the inherent dangers of nuclear weapons and those who identify benefits in nuclear technology and the production of nuclear energy. This balance has facilitated both NPT and IAEA to carve out a unique position and enabled it to obtain almost universality in its application.

As with CWC, the history of NPT therefore suggests that, by taking steps to build a balanced partnership with all those concerned by the subject matter, it is possible to foster widespread support for sensitive and divisive matters and to keep both sides content with the outcome. In this regard, when approaching the issue of autonomous robotics, there may be merit in carefully looking at NPT and considering a right of states to pursue scientific and technological advancements. As a means to striking a balance between those in the

domestic and industrial robotics world and those against the military application of autonomous robotics, this right could be set against a mechanism for transparency and openness, akin to Article III of NPT and the IAEA safeguards. The successes of both CWC and NPT clearly indicate that, ultimately, if the system of governance is to be effective, the approach must be equitable.

Wide awake and walking towards the future

Society is on a threshold with regard to autonomous robotics and the coming years are certain to be definitive as the robotics industry refines its capacity to produce robots and these robots become increasingly integrated into mankind's daily life. Some believe that, with the leaps and bounds being made in technology every day, we are quickly approaching the futurist theory known as the 'Singularity' - a period proposed by computer scientist and author Ray Kurzweil "during which the pace of technological change will be so rapid, its impact so deep that technology appears to be extending at infinite speed." (25) At this point in time, Kurzweil believes, artificial intelligence will surpass human intelligence. Proponents of this theory believe that these changes are inevitable and that the 'Singularity' will radically alter the future of society. It is of course difficult to say, with any degree of certainty, whether this will or will not be the case and, if it is, when it will occur. What we should be able to agree on at this point in time is how very much unprepared we are for whatever the future may hold. The good news is, however, that autonomous robotics — or at least killer robots for now — are finally beginning to appear on the international radar as something for consideration.

In November 2014, another round of discussions at the level of the United Nations will be held in Geneva within the scope of the Convention on Certain Conventional Weapons. While this meeting will most probably not enter into any specific details of a ban, or a Convention setting out a ban, it is the second such high profile event on the subject matter and will be a critical venue for future discussion. It is hoped that the present discourse will serve to prompt an earnest and comprehensive discussions of autonomous robotics, taking into account the importance of building upon the lessons learned from other international instruments. Thus far, remarkable results have been achieved under CWC and NPT facilitated by the almost universal level of support for each instrument: a fact that must be given due recognition for any future success with regard to autonomous robotics. It may yet be some time before a true system of governance of autonomous robotics is in place, but we can already start to visualise this system by fostering a comprehensive and balanced approach to the world of autonomous robotics.

Given that Asimov's fictional laws of robotics are still all we have to show for a system on autonomous robotics, the time to stop "sleepwalking" is now. Without over-indulging in speculation, we are, as author James Barrat, observes, only one '9/11' away from learning another set of excruciating lessons. The only difference with other disasters, he notes, is that with autonomous robotics "they're not like airplane disasters, nuclear disasters, or any other kind of technology disasters with the possible exception of nanotechnology. That's because there's a high probability we won't recover from the first one." (26) On the 30th anniversary of when the world first met Arnold Schwarzenegger's Terminator, now would not only be a timely, but almost the perfect moment, to take the first steps for the future system for the governance of autonomous robots. (27)

The authors:

Irakli Beridze is a Senior Strategy and Policy Advisor at UNICRI, previously serving as Special Projects Officer at the Organisation for the Prohibition of Chemical Weapons (OPCW). During his 15 years of service at a number of international organizations, he was responsible for various programmes and projects in the area of international security, CBRN, non-proliferation, disarmament, counter-terrorism, security governance and emerging technologies. He is a member of international task forces and working groups, and has undertaken missions in various countries of the world. He has degrees in International law, International relations and Political Science. In 2013 he received recognition on the occasion of the awarding of the Nobel Peace Prize to the Organisation for the Prohibition of Chemical Weapons.

Odhran James McCarthy is as a Project Officer at UNICRI on the CBRN Risk Mitigation and Security Governance Programme and is involved in the implementation of the EU CBRN Risk Mitigation Centres of

Excellence Initiative and the EU initiative on Strengthening Bio-Safety and Bio-Security Capabilities in the South Caucasus and Central Asian Countries. He has a background in international law and the law of international organisations and is particularly interested in the safety and security implication of autonomous robotics and other emerging technologies.

1 [http://www.unog.ch/80256EDD006B8954/%28httpAssets%29/12688EA8507C375BC1257CD70065815B/\\$file/Brazil+MX+LAWS.pdf](http://www.unog.ch/80256EDD006B8954/%28httpAssets%29/12688EA8507C375BC1257CD70065815B/$file/Brazil+MX+LAWS.pdf)

2 <http://www.irobot.com/us/learn/home/roomba.aspx>

3 <http://www.robomow.com/en-USA/>

4 <http://www.myjibo.com/>

5 <http://www.nbcnews.com/watch/nbc-news/pie-in-the-sky-russian-pizza-chain-delivers-by-drone-287734851748>

6 U.S. Joint Forces Command, Unmanned Effects (UFX): Taking the Human Out of the Loop, Rapid Assessments Process (RAP) Report 03-10, 2003, p.5.

7 Max Booth considers this future unlikely and suggests that “machines will only be called upon to perform work that is dull, dirty, or dangerous” - Boot, Max, War made War: Technology, Warfare, and the Course of History, 1500 to Today, Gotham Books, 2006 p.442.

8 Sharkey, Noel, 2004: Big robot is watching you: Report on the future of robots for policing, surveillance and security, University of Sheffield, 2008, p.4.

9 Asimov, Isaac, I, Robot, Dobson Books Ltd, 1967. Asimov later introduced a fourth or “Zeroth Law” stating that “A robot may not harm humanity, or, by inaction, allow humanity to come to harm.”

10 <http://www.stopkillerrobots.org/>

11 UN Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Professor Christof Heyns - http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf

12 <http://www.motherjones.com/politics/2013/05/campaign-stop-killer-robots-military-drones>

13 <http://www.washingtonpost.com/wp-dyn/content/article/2007/06/28/AR2007062801338.html>

14 <http://www.reuters.com/article/2013/12/16/us-google-robots-idUSBRE9BF17X20131216>
http://www.nytimes.com/2013/12/04/technology/google-puts-money-on-robots-using-the-man-behind-android.html?pagewanted=all&_r=1&#h

15 <https://www.dvidshub.net/news/135952/meeting-ls3-marines-experiment-with-military-robotics#.U-s5W2PN225>
<http://www.bbc.com/news/technology-25493584>

16 The Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on their Destruction, which is also more simply known as the Biological Weapons Convention (BWC), is a possible third Convention that could be reviewed for best practices. Indeed, many of the lessons learned from NPT and CWC that are detailed below have analogous lesson in BWC. The most notable here would be Article X that encourages peaceful uses of biological science and technology. However, for the sake of brevity this discussion will focus on two of the more critically renowned instruments in non-proliferation, namely CWC and NPT.

17 <http://www.opcw.org/news-publications/publications/facts-and-figures/>

18 http://www.opcw.org/index.php?eID=dam_frontend_push&docID=6357

19 OPCW Non Proliferation Seminar, René van Sloten, Executive Director, Industrial Policy, European Chemical Industry Council (CEPIC), 11 April 2011 http://www.opcw.org/fileadmin/OPCW/events/2011/NPS/papers/opening/Dr_Rene_van_Sloten_CEPIC.pdf

20 Ibid.

21 Scheduled chemicals are the chemicals and their precursors that are listed in Schedules 1, 2 and 3 in the Annex to CWC. The General Purpose Criterion in essence acknowledges that the industry is not static and therefore the list of scheduled chemicals can never be truly exhaustive. As such, the criterion requires that any other chemical used for purposes prohibited by the Convention equally fall within the remit of the Convention. This criterion is considered to be a key tool that allows the Convention to continue to be applicable in spite of developments in the chemical industry.

22 Helsinki Chemical Forum, "Chemical Industry and Chemical Weapons Convention", Ambassador of Japan, Minoru Shibuya, Chairman of the Conference of the States Parties to The Chemical Weapons Convention, 29 May 2009.
http://www.helsinkicf.eu/wwwcem/cem/program/Materials/Finland_Chemical_Forum_Statement1.pdf

23 Ambassador Ahmet Üzümcü, Director-General of the OPCW, Welcome remarks by the Director General, Informal meeting between States Parties and the chemical industry, 24 September 2012
http://www.opcw.org/fileadmin/OPCW/ODG/uzumcu/DG_Statement_-_Meeting_with_industry_24_Sept_2012.pdf

24 <http://www.iaea.org/Publications/Documents/Infcircs/Others/infcirc140.pdf>

25 Kurzweil, Ray, The Singularity Is Near. When Humans Transcend Biology. Viking, 2005. pp 24.

26 Barrat, James, Our Final invention, Artificial Intelligence and the End of Human Era, St. Martin's Press, 2013, pp. 27-28.

27 Terminator was released in the United States on 26 October 1984.