Please Cite This Report as:

D. E. Koditschek, L. M. Titus, and J. Hamilton, "Report on the ICRA'22 Workshop on Lethal Autonomous Weapons," University of Pennsylvania, Philadelphia, PA, GRASP Lab Report, Dec. 2022. https://repository.upenn.edu/ese_reports/27/

Report on the ICRA'22 Workshop on Lethal Autonomous Weapons[1]
Expanded Version for the Penn GRASP Faculty by D. E. Koditschek
of a Report Submitted to the Perry World House by the Workshop Organizers

Brief Description

The full-day hybrid (in-person and remote) workshop at the Institute of Electrical and Electronics Engineers (IEEE) International Conference on Robotics and Automation (ICRA) took place on Friday, May 27th, 2022 in Philadelphia, PA. It presented an interdisciplinary lineup of scholars and interactive sessions to help robotics researchers develop a disciplinary response to the challenges surrounding legal and ethical governance in design and use of Lethal Autonomous Weapons Systems (LAWS).

The workshop organizers were: Lisa Miracchi Titus (Organizer and Chair), Michael C. Horowitz (Co-Organizer), Daniel E. Koditschek (Co-Organizer), Victoria Edwards, Walker Gosrich, Jesse Hamilton, Ariella Mansfield, and Shane Rozen-Levy.

After outlining the workshop goals and speaker profiles, this report provides a **Synopsis of Workshop Proceedings** including summary abstracts of the three workshop sessions and speakers' talks as well as links to the publicly available recordings of each talk provided by the TechEthics Channel on IEEE.tv. It concludes with brief notes on the **Participation** by workshop attendees and their experience as observed in breakout sessions and subsequent questionnaire responses.

Outline of Workshop Goals

1. Facilitate interdisciplinary communication and connections between roboticists and researchers in different fields on the subject of LAWS, to better understand differing perspectives and opinions as well as opportunities for progress.
2. Advance our understanding of technical and ethical issues related to LAWS, and document these insights as workshop proceedings in a suitable archival venue.
3. Create the foundation for lasting working groups to propose well-articulated positions on the regulation of autonomy in robotic weapons systems, with the long term goal of proposing to the RAS Research and Practice Ethics Committee a path toward what robotics as a discipline can contribute to the governance of AWS.

Overview of Workshop Speakers

- Ruzena Bajcsy, GRASP Lab Founder,Professor Emerita, Electrical Engineering, and Computer Science, UC Berkeley, Member, Robotics and Automation Research and Practice Ethics Committee.

Please Cite This Report as:

- [Ariel Conn](), heads the IEEE-SA Research Group on Issues of Autonomy and AI for Defense Systems.
- [Claire Finkelstein](), Algernon Biddle Professor of Law and Professor of Philosophy, University of Pennsylvania.
- [Denise Garcia](), Associate Professor of Political Science and International Affairs, Northeastern University, co-author of IEEE SA Report.
- [Ryan Gariepy](), CTO, Clearpath Robotics/OTTO Motors, Board of Directors, Open Source Robotics Foundation, and Co-Chair, Canadian Robotics Council.
- [COL Christopher Korpela](), Associate Professor, Robotics Research Center Director, U.S. Military Academy, West Point.
- [Jonathan Moreno](), David & Lyn Silfen University Professor, Professor of Medical Ethics & Health Policy and of History & Sociology of Science, Penn Integrates Knowledge (PIK) Professor, Professor of Philosophy, University of Pennsylvania.
- [Brendan Schulman](), Vice President of Policy & Government Relations at Boston Dynamics.
- [Lisa Miracchi Titus](), Associate Professor of Philosophy and GRASP Affiliate Faculty, University of Pennsylvania.

Synopsis of Workshop Proceedings

Session 1: Complexity of the Ethical Issues for LAWS
In what ways are discussions about the ethics of LAWS continuous with historical discussions about the ethics of increasingly autonomous weapons? Do LAWS pose a distinctive, in-principle problem to their ethical application? What complexities arise in the development of autonomous weapons in a global, multi-cultural context?
More information on Session 1 can be found [here]().

["Autonomous Armed Robots and the Principle of Distinction: Does Robotic Killing Violate the Laws of War]()?" – Claire Finkelstein
This talk explored the extent to which the Law of Armed Conflict (LOAC) already applies to extant and future weaponized robots and other new technologies of war. Ethical and legal problems were never properly addressed throughout the vast increase over the past decade's use of lethal semi-autonomous technology, leaving a legacy of complicated questions. Between jus ad bello vs jus in bello (necessity, proportionality, distinction) attention arguably should focus on the in-bello principle of "distinction" (e.g., in contrast to the case of nuclear technolgy for which the key issue would likely be "proportionality"). This entails distinctions to be made along a complicated spectrum of participants from armed combatant (always targetable) to unarmed civilian (never targetable). Since humans have such difficulty in establishing these distinctions, how might robots ever be able to do so? Arguably, machines would engage in top down reasoning whereas humans engage in bottom up (analogical) reasoning. But do we even want robots to engage in human-style moral reasoning? Moreover, will we lose our grip on principles of responsibility (roughly analogous to keeping wild animals domestically and other contexts)? Importantly, in addition to use restrictions (typically too little too late) we must also consider restrictions on development of such techologies.

Please Cite This Report as:

D. E. Koditschek, L. M. Titus, and J. Hamilton, "Report on the ICRA'22 Workshop on Lethal Autonomous Weapons," University of Pennsylvania, Philadelphia, PA, GRASP Lab Report, Dec. 2022. https://repository.upenn.edu/ese_reports/27/

"International Legal and Diplomatic Challenges Posed by Autonomous Weapons and AI-Assisted Systems" – Denise Garcia

This talk focused on the impact of LAWS on the nature and status of international law. Substantially strengthened in recent decades, these humanitarian advances will be threatened if we don't properly regulate autonomous weapons because they challenge the assignment of responsibility and because they rely upon technologies that are becoming ubiquitous. The current framework of international humanitarian law (IHL) will not suffice to cover autonomous weapons but such weaspons have not yet been deployed – hence we don't yet have victims with legal standing to challenge them.
Despite much discussion of LAWS over the past decade by the Geneva Convention Group of Government Experts (GGE) on Certain Conventional Weapons (CCW), these talks are presently badly boggd down and armed nations have further obfuscated progress. At best, there is a rough consensus on the need for "meaningful human control" that would install certain prohibitions and impose additional regulations. This is urgent to move forward within an international framework.

"Defining the technical challenges of autonomous weapons systems to address ethical and policy challenges" – Ariel Conn

This talk focused on the origin and nature of the IEEE Standards Association Report on challenges in the development, use and governance of autonomous weapons systems. The many different parties to the Geneva Convention GGE talks on CCW roughly agree on the need for meaningful human control but it's very hard to translate such an abstract phrase into actionable terms – much less, executable code. Hence the reported challenges (10 categories summarize over 60 specific challenges) address the need for disciplinary expertise in advising both the technical and regulatory communities. The challenge of lacking a common language arises from the huge technological gulf between between today's advanced drones on the one hand and general AI on the other that is obscured by the many different parties using the same terminology across this gulf. The challenge of insuring meaningful human control (notwithstanding the apparent unanimity concerning its centrality across all the GGE) similarly arises from the difficulty in pinning down what that term means operationally.

Session 2: Defining and Refining Challenges for Robotics

With this context at the forefront, we address issues around defining ethical requirements for development and testing. Human control and responsibility are clearly required for the ethical use of AWS, especially LAWS, but how should we understand these requirements? How does the diversity of human behavior affect these requirements? What do roboticists need to establish about the behavior of their systems in order to make human control and responsibility possible in real military contexts?
More information on Session 2 can be found here.

"The Neuroethics of LAWS from Neural Networks to Robotics and Back Again" – Jonathan Moreno

This talk explored aspects of LAWS that intersect with issues surrounding human machine interaction in the military context, reflecting the speaker's longstanding concerns, presently represented by a DoD funded study on the ethics of human war-fighters' trust in LAWS. Observing the research transition from "good old fashioned AI" to Neural Network-based AI, there seems to be occurring a convergence of technologies between AI/robotics, neuroscience and brain-machine interfaces. A noteworthy token of

these developments is reflected in the growing multitude of US DoD (along with many other nations' militaries') sponsorship of neuroscience projects with an emphasis on neuro-interfaces and cognitive enhancement (including interfacing). Analogous to the very different thresholds for the use of nuclear weapons across the globe (e.g. Russia has a much lower announced trigger for the introduction of nuclear weapons in combat) important questions arise in considering whether different nations will enforce different triggers relative to the use and ethics of these new technologies.

"Permissible Uncertainty and Meaningful Human Control" – Lisa Miracchi Titus

This talk argued that, given the differences between AWS and genuinely intelligent systems, we should focus on understanding what it takes for the humans involved to be responsible for the behavior of such systems. This involves developing a conception of what meaningful human control (MHC) amounts to for such systems. Titus proposes understanding the concept of MHC in terms of *permissible uncertainty* - - that is, the uncertainty that it is permissible for a human agent to have given the role(s) she plays regarding such systems. This helps us productively understand continuities between the use of AWS and the deployment of human agents in combat as well as some of the difficulties distinctive to responsibly using AWS systems. In particular, it can help us to specify needed areas of communication between ethicists and roboticists, so that we can clarify what conditions must be met at design, development, and testing stages in order for military personnel to be in a position to make responsible decisions about the use of AWS.

"Why Is No One Banning Killer Robots?" – Ryan Gariepy

This talk addressed the question of why it has proven so hard to advance the international discussion on where robots do not belong – particularly in regard to their weaponization. The key insight is that international discussions are intrinsically political, thus often national and commercial interests militate against the development of constraints. A central unresolved problem in these discussions remans the problem of defining meaningful human control. A suggested guiding principle for any effective regulation is to aim for prescriptive rather than prohibitive expressions of constraints since it is easier to ask for proof of standardized characteristics than to prove their absence. Given the widespread and extensive use of semi-autonomous lethal weapons systems there is a growing need to address the vast gray area and need to distinguish systems that are already deployed from those that are under development, as well as from those that loom in the near future. This must be advanced by providing and analyzing specific examples of in-use or in-development cases. It is also crucial to begin recognizing the increasingly ubiquitous nature of these systems (e.g. by police in civilian spaces, e.g. by civilian property owners) and hence, that these discussions are likely to increase in complexity. Finally, more roboticists are needed to keep this international discussion on track, where they should do what they can to insure that robots increase rather than decrease social well-being.

Session 3: Shaping a Disciplinary Response in Robotics
In light of the articulation of ethical questions and challenges in Session 1, and the clarification of issues around development and testing in Session 2, the last panel asks robotics researchers, military, and industry experts for guidance in shaping a disciplinary response in robotics to ethical issues of LAWS. How should sensitivity to these issues shape research practice? Education and training? What possibilities might there be to shift funding incentives to

promote the thorough integration of ethical concerns in engineering practice from conceptualization through application?
More information on Session 3 can be found here.

"Weaponized Robots: Understanding the Spectrum of Risk and Fear" – Brendan Schulman

This talk assessed the nature of challenges and problems associated with weaponized robots. In this view, the central concerns arise from their potential to undermine public safety and public trust. Autonomy is not the only concern given the steady increase in mobility that advances the efficacy of remote operation, raising questions about unseen operators' accountability, ability to communicate intent, as well as their removal from any immediate adverse consequences attending their potential aggression or misdeeds. The robot drone industry several years ago organized as a group to develop an appropriate set of guidelines and context that, for example, has led to the FAA requirement for a uniquely identifying broadcast serial number associated with every machine. In the absence of any equivalent set of guidelines for ground robots, Boston Dynamics has chosen to "bake into" its Spot quadruped product line terms and conditions of purchase that preclude weaponization on pain of lost warranty and access to further upates. Yet legs don't seem to be the problem in and of themselves. Rather, there seems to be an intricate mix of perceived risks and fears that must be untangled along various axes of potential concern. Among these is surely the quality of a machine's mobility ranging from the most futuristic terminator-style agility through the terrain capabilities of present machines, declining with confined indoor operation or complete immobility. Similarly, there is a spectrum of lethality to be considered ranging, for example, from sound, through taser, tear gas, BB-gun, knife, gun, machine gun, and, ultimately, explosives. The social context is also likely in play, for example, spanning an axis of increased concern about operator identity ranging from a national military, to a civilian government, to a known individual, through an unknown criminal. Of course, the degree of autonomy remains foremost, ranging from remote controlled platform to one possessing some autonomous mobility through to advanced autonomy on the scale of present-day science fiction.

"Autonomous Rules of Engagement Escalation in Human-Robot Teams" – COL Christopher Korpela

This talk reviewed the nature of US DoD efforts to govern ethically the deployment and use of LAWS, given from the perspective of a military researcher focused on the ethics of human-machine teaming who has also served as a US representative to Geneva's GGE discussions on CCW. The overarching observation is the need for a "whole-of-society" effort in representing the great diversity of views on these topics. Beyond its general (US DoD 3000.09) guidelines, US has been an avid participant in the GGE CCW addressing the challenges of definitions, the crucial issues of operational context (e.g., urban vs. rural; spatiotemporal constraints) and exploring use-cases. Despite the many frustrations of this setting, it represents a critical international discussion where everyone needs to be heard. The major points of consensus are to promote existing laws of armed conflict and emphasize human responsibility via the chain of command. Major outstanding challenges remain concerning how to define and pursue bad actors who misuse these technologies (countering war crimes) as well as how to guarantee predictability and reliability in the technologies themselves. However, a consideration of the benefits of LAWS helps

motivate DoD's keen intent on its development (albeit, particularly, while ensuring its compliance with ethical principles and compliance with international humanitarian law). Precision guided weapons advance the principle of proportionality by permitting smaller submunitions to deliver smaller charge with greater specificity. Autonomous agents with increased sensing and computational capabilities can take much greater risks to get much better information, advancing the principles of distinguishability.

"Roboticists' Social Responsibility" – Ruzena Bajcsy

This talk addressed robotics researchers' responsibility for the applications of the technologies they invent and, in particular, for the development and use of weaponized robots. The speaker, having personally witnessed both Hitler's and Stalin's totalitarian states, emphasized her motivation as arising from the crucial importance of defending democracy by all possible means. Coming out of central Europe which was a focus of much of the 20th century strife, Capek originally invented the word "robot" to express the slavish nature of Hitler's followers. Roboticists might compare their present situation to that of the physicists who invented nuclear bombs. Two years after Hiroshima, they convened the Pugwash conference in 1947. This event was in large measure motivated by a manifesto authored by Bertrand Russell and Albert Einstein which called upon worldwide scientists of all political persuasions to discuss the implications of the nuclear bomb. Sakharov's globally witnessed mistreatment and exile consequent upon his participation in these efforts epitomize the resistance of dictators to ethical discussions. Today, a similar responsibility faces all robotics researchers. It cannot suffice to simply reject military support: the long history of science and technology advancements reveals their intimate association with defense funding. This record of militarized science and technology (including some history of abuse) insures that even those who choose to reject defense funding are still complicit in the process and implications of the military technology project. There is no escape from this responsibility. Since roboticists know best the limitations as well as capabilities of our research, we must be very clear about what doesn't work and what are the underlying assumptions along with the present and likely future weaknesses. Particularly, we must inform our sponsors: on the one hand it is crucial for roboticists to not oversell; similarly, sponsors have the responsibility to ask careful appropriate questions. Robotics researchers have still more responsibility.

Participation

Overall, the workshop was well-attended. Out of some 200 pre-registered participants, the workshop attracted roughly 60 in-person and 15 remote attendees. The contributed session included five posters on a variety of LAWS-related topics from a diverse set of researchers and doctoral students.

Additionally, the workshop hosted a poster session aimed at recruiting interdisciplinary researchers to partner with robotics researchers in developing a disciplinary response to the challenges surrounding legal and ethical governance in the design and use of lethal autonomous weapons systems (LAWS) and related technologies. More information on the poster session can be found here.

"Artificial Corporations: Moral Norms in Artificial Deep Learning Systems" – Sophia Wushanely, University of Michigan

"By Air or by Land: How Locomotion Methods Dictate Drone Ethics" – James Zhu and Aaron M. Johnson, Carnegie Mellon University

"Crafting Quality Law And Policy For Robotics" – Sogol Balali, Ross T. Sowell, Ruth West, Cindy Grimm

"Evaluating the Use of Lethal Autonomous Weapons in Checkpoints" – Russel Perkins and Paul Robinette, University of Massachusetts (Lowell)

"We must regulate law enforcement's use of robots" – Sonia Roberts (Northeastern University), Shane Rozen-Levy (University of Pennsylvania), and Matthew Malencia (University of Pennsylvania)

Participant Response

Realtime in-situ notes of the various breakout sessions taken by the organizers suggest that the overarching goals of the workshop were met. The depth and nuance of many discussions suggests a successful encounter of roboticists with ideas from the other disciplines represented by the various speakers' presentations. Moreover, grouping similar threads of ideas across the various sessions reveals a substantial overlap of participants' concerns with some of the salient issues identified in the IEEE Challenges report concerning the difficulty of predictability and explainability and consequent obstacles to verification and validation. At the same time, a common thread across breakouts distinct from the concerns of the IEEE Challenges report is the need for the field to incentivize honest reporting of technical limitations alongside the need for more systematic ethics education. Arguably, both sets of issues – the difficulties in developing standards for verification and validation as well insufficient ethical education and research reporting obligations – arise in part from a lack of disciplinary foundations, limiting the present capacity of robotics to respond to the technical and ethical challenges that its growing technological successes necessitate.

The close of the day questionnaire yielded some thirty positive responses (seemingly half from students, a quarter from faculty and the remainder from government or corporate employees) to the invitation to form a ongoing groups addressing the workshop's topics, with a preponderance of interest in a robotics research focused working group.