Dear Mr. President,

We, the undersigned scientists and engineers, urge that your administration undertake an immediate and intensive effort to formulate and secure an international treaty restricting the development, deployment, and use of lethal autonomous weapons systems – robotic weapons that can locate, select, and attack human targets without human intervention. Whereas current US policy, laid out in the remarkably farsighted DoD Directive 3000.09, requires "appropriate levels of human judgment over the use of force" and specifically disallows autonomous selection of human targets even in defensive settings, a treaty would extend and strengthen these restrictions in the international sphere. This step is an urgent priority given the current rate of technological progress. Absent a treaty, we foresee the prospect of an arms race with negative outcomes for both humanitarian and strategic concerns: in particular, it may lead to a new class of "scalable" weapons of mass destruction – weapons that even small groups could use to attack large populations. Rather than constituting a "third offset" to maintain US military dominance, these developments would instead pose a threat to US and international security.[1]

Legal and humanitarian considerations

UN Special Rapporteur Christof Heyns, Human Rights Watch, the International Committee of the Red Cross, and other experts have expressed concerns about the ability of autonomous weapons to comply with provisions of the laws of armed conflict regarding military necessity, proportionality, and discrimination between combatants and civilians. We do not believe that compliance is feasible at present or in the near future; it requires that machines make subjective and situational judgments that are considerably more difficult than the relatively simple tasks of searching for and engaging potential targets. Even if compliance becomes technically possible, there is of course no guarantee that all parties would use autonomous weapons in legally compliant ways.

Delegating to a machine the decision over the life or death of a human being also raises a fundamental moral question. The Martens Clause of the Geneva Conventions declares that, "The human person remains under the protection of the principles of humanity and the dictates of public conscience." In this regard, Germany has stated that it "will not accept that the decision over life and death is taken solely by an autonomous system" while Japan "has no plan to develop robots with humans out of the loop, which may be capable of committing murder."[2] BAE Systems, the world's second-largest defense contractor, has asserted that it has no intention of developing autonomous weapons, stating that the removal of the human from the loop is "fundamentally wrong."[3] At present, the broader public has little awareness of the state of technology and the near-term possibilities, but this will presumably change if and when video footage of robots hunting down and killing unarmed children starts to emerge from distant urban battlefields. At that point – regardless of the identity of the ultimate perpetrators – the dictates of public conscience will be very clear but it may be too late to follow them.

---

[1] A letter from British scientists, making similar arguments, is being sent to Prime Minister David Cameron.

[2] Statements by the respective ambassadors to the CCW meeting in Geneva, April 2015.

[3] Statement by Sir Roger Carr, BAE chairman, at the World Economic Forum, January 21, 2016; https://www.youtube.com/watch?v=opZR7vLhXVg.

Strategic considerations

The component technologies for autonomous weapons, including automated decision making, computer vision, robotics, control systems, and precision manufacturing, have reached the point where fully autonomous weapons are currently feasible for many aerial and naval missions and may soon be feasible for urban warfare. An arms race in autonomous weaponry will lead inevitably to low-cost, mass-produced devices such as flying micro-robots able to hunt for and eliminate humans inside buildings. Such devices will form a new, scalable class of weapons of mass destruction with destabilizing properties similar to those of biological weapons. Their scalability is tied intrinsically to their autonomy: once available in large numbers on the arms market, they can be acquired, managed, and launched in the millions with few personnel and almost no infrastructure. Thus, they tip the balance of power away from legitimate states and towards terrorists, criminal organizations, and other non-state actors. In addition, they are ideal tools for repression by authoritarian regimes, requiring no bribes or special privileges and having no conscience.

The considerations of the preceding paragraph apply principally to weapons designed for ground warfare and anti-personnel operations, and are less relevant for naval and aerial combat. It is still the case, however, that to entrust a significant portion of our defense capability in any sphere to autonomous systems is to court instability and risk strategic surprise. Autonomous weapons in conflict with other autonomous weapons must adapt their behavior quickly, or else their predictability leads to defeat. This adaptability is necessary but makes autonomous weapons intrinsically unpredictable and hence difficult to control. Moreover, the strategic balance between robot-armed countries can change overnight thanks to software updates or cybersecurity penetration, leading to potentially incorrect perceptions of security or strategic superiority. Finally, the possibility of an accidental war – a military "flash crash" involving spiraling and unpredictable high-speed interactions among competing algorithms – cannot be discounted.[4] Thus, while there are many ways in which AI and related technologies can contribute to the maintenance of US strategic superiority, the development of fully autonomous weapons does not appear to be one of them.

We believe it is possible to design a treaty banning lethal autonomous weapons that will avoid these risks and forestall the large-scale manufacturing that would result in the wide dissemination of these scalable weapons. Although limiting proliferation of these technologies comes with unique challenges, experience with the Chemical Weapons Convention suggests that, with industry cooperation, the residual threat from the diversion of dual-use technology into "home-made" weapons may remain manageable. Moreover, defensive anti-missile systems and anti-robot countermeasures could and should remain in place.

Conclusion

Your leadership in this matter is crucial for the current treaty discussions to move forward. We see a window of opportunity, albeit one that is short due to rapid advances in technology. While we do not underestimate the difficulty of the negotiations, we are confident that a treaty can be concluded that maintains and improves US national security and would become an important legacy of your presidency.

---

[4] A recent report from the Center for a New American Security, "Autonomous Weapons and Operational Risk," makes many of the same points.

Signatories

Stuart Russell, Smith-Zadeh Professor of Computer Science, University of California, Berkeley
Tom Dietterich, Distinguished Professor, Oregon State University; President of AAAI
Nils J. Nilsson, Kumagai Professor of Engineering, Emeritus, Stanford University; Past President of AAAI
Barbara J. Grosz, Higgins Professor of Natural Sciences, Harvard University; Past President AAAI; former Chair of IJCAI Board of Trustees; National Academy of Engineering
Tom Mitchell, Fredkin University Professor, Carnegie-Mellon University; Past President of AAAI; National Academy of Engineering
Eric Horvitz, Managing Director, Microsoft Research; Past President of AAAI; National Academy of Engineering
Martha E. Pollack, Professor of Computer Science and Provost, University of Michigan; Past President of AAAI
Henry Kautz, Professor of Computer Science, University of Rochester; Past President of AAAI
Ronald J. Brachman, Former Chief Scientist, Yahoo!; Past President of AAAI
C. Raymond Perrault, Director, AI Center, SRI International; Past President of IJCAI
Donald Knuth, Professor Emeritus of The Art of Computer Programming, Stanford University; National Academy of Sciences; National Medal of Science
Noam Chomsky, Institute Professor emeritus, MIT; National Academy of Sciences
Elon Musk, CEO, SpaceX; CEO, Tesla Motors; Chairman, Solar City; Co-Chairman, OpenAI
Steve Wozniak, Co-founder, Apple Inc.
Frank Wilczek, Herman Feshbach Professor of Physics, MIT; Nobel Laureate
George Church, Professor of Genetics, Harvard University; National Academy of Sciences; National Academy of Engineering
Peter Norvig, Research Director, Google
Donald Norman, University of California, San Diego; National Academy of Engineering
Michael Jordan, Professor of Computer Science and Statistics, University of California, Berkeley; National Academy of Sciences; National Academy of Engineering
Daniel C. Dennett, University Professor and Fletcher Professor of Philosophy, Tufts University
Ruzena Bajcsy, NEC Professor of Electrical Engineering and Computer Sciences, University of California, Berkeley; National Academy of Engineering
George A. Bekey, Professor of Computer Science (Emeritus), Univ. of Southern California; National Academy of Engineering
Oren Etzioni, CEO, Allen Institute for AI
Corinna Lathan, CEO, AnthroTronix; Chair, World Economic Forum Council on AI and Robotics
Bart Selman, Professor of Computer Science, Cornell University
Joseph Y. Halpern, Professor of Computer Science, Cornell University; American Academy of Arts and Sciences
David C. Parkes, Professor and Area Dean for Computer Science, Harvard University,
Berthold K.P. Horn, Professor of Computer Science and Engineering, MIT; National Academy of Engineering
Jean-Claude Latombe, Professor (Emeritus) of Computer Science, Stanford University
Max Tegmark, Professor of Physics, MIT
Steven Minton, President, InferLink Corporation
Michael Wellman, Professor of Computer Science & Engineering, University of Michigan
Leslie Pack Kaelbling, Panasonic Professor of Computer Science and Engineering, MIT
John E Laird, Professor of Computer Science and Engineering, University of Michigan

Tuomas Sandholm, Professor of Computer Science, Carnegie Mellon University; CEO, Optimized Markets
Richard E. Korf, Professor of Computer Science, University of California, Los Angeles
Zoubin Ghahramani, Professor of Information Engineering, University of Cambridge; Fellow of the Royal Society
Jeff Dean, Senior Fellow, Google, Inc.; National Academy of Engineering
J. Strother Moore, Admiral B.R. Inman Centennial Chair Emeritus of Computing Theory, University of Texas at Austin; National Academy of Engineering
Barbara Simons, IBM Research (retired); Past President of ACM
William Oldham, Professor Emeritus of Electrical Engineering and Computer Sciences, University of California, Berkeley; National Academy of Engineering