# 2023 Executive Order on Trustworthy AI Misses Issues of Autonomy and AI Multi-Threat Challenges

Chris C. Demchak, Ph.D.
Sam J. Tangredi, Ph.D.

On October 30, 2023, the Administration released Executive Order (EO) 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (AI). For several reasons this EO lays a light hand on the Department of Defense (DoD).

First, the DoD had and was already implementing "responsible AI" safeguards for defense programs.[1]

Second, as the federal agency most likely interested in "killer robots," the DoD for years has firmly embedded an ironclad policy to ensure that "[t]here is always a human always responsible for the use of force" – the so-called "human-in-the-loop (HITL)."[2]

Third, peer adversaries to the U.S. have yet to openly demonstrate autonomous capabilities that could force the military to publicly reconsider its ethical principles, even though we do not naively expect adversaries to observe ethical constraints in combat.

Fourth, non-DoD U.S. government agencies are a step behind the military in setting, much less enforcing, policies on responsible AI.[3]

Fifth, a public debate over AI for some time has been in progress among political and IT capital goods elites over a rising concern that uncontrolled commercial and government use of AI is spreading too rapidly into all private and consequential aspects of citizens' lives – areas in which DoD generally plays a limited role.

With engineering, economics, and comparative complex organization theory/political science degrees, **Dr. Chris C. Demchak** is the RDML Grace M. Hopper Chair of Cyber Security and the Senior Cyber Scholar, Cyber and Innovation Policy Institute, U.S. Naval War College. Her research and many publications cyberspace as a globally shared insecure complex 'substrate' take a systemic approach to emergent structures, comparative institutional evolution, adversaries' use of systemic cybered tools, virtual worlds/gaming for operationalized organizational learning, and designing systemic resilience through collective cyber defense against imposed surprise.

Sixth, while the President cannot enforce these principles on the nation's commercial vendors, he can dictate policy and procedure to federal agencies as to how they pursue their AI, and through their purchases, perhaps indirectly influence the wider national AI market.

While these are all solid justifications for the EO's limited attention to the DoD, it is worth exploring how consideration of two related and forthcoming realities might have made a difference if included – autonomy and dual uses of AI.

### *Contents of the EO*

The Executive Order weighs in at 36 pages. Its stated purpose is to reduce the potential for "irresponsible use" of AI which it defines as "exacerbate[ing] societal harms such as fraud, discrimination, bias, and disinformation; displac[ing] and disempower[ing] workers; stifl[ing] competition; and pos[ing] risks to national security."[4] As can be detected right away, its examination is more on "societal harms," than extensive risks to national security. However, it does make substantial reference to "dual-use" technology and necessary cooperation between other departments and DoD to establish further policies on related risks.

The EO gives directives to government departments and agencies in a seemingly firm and decisive manner. As an example: "The Secretary of Labor shall, within 180 days of the date of this order and in consultation with other agencies and with outside entities, including labor unions and workers, as the Secretary of Labor deems appropriate, develop and publish principles and best practices for employers that could be used to mitigate AI's potential harms to employees' well-being and maximize its potential benefits."[5]

However, this apparent decisiveness is mitigated by the fact that the EO applies *only* to government agencies. It has no authority over commercial or non-government organizations. Thus, to use the Labor Department

**Dr. Sam J. Tangredi** is the Leidos Chair of Future Warfare Studies and professor of national, naval and maritime strategy at the U.S. Naval War College. A retired Navy captain and surface warfare officer and strategic planner, he has published seven books, over 200 journal articles and book chapters, and numerous reports for government and academic organizations. His latest book (co-authored with George Galdorisi) is *Algorithms of Armageddon: The Impact of Artificial Intelligence on Future Wars* (Annapolis, MD: Naval Institute Press, 2024). He held command at sea and directed strategic planning organizations.

example, there is nothing to compel "outside entities" to cooperate with directive activities if they chose not to cooperate. What if labor unions and worker have no desire to be in "consultation" with the government? Can the practical objectives of the EO—as opposed to the churning out of published "best practices" and other white papers—be fulfilled if the non-government sector ignores it? This is particularly acute concerning corporations developing AI which have every desire to take every legal action they can to turn a profit, oppose government regulation, and keep control over their intellectual property.

AI development is dominated by commercial firms for commercial purposes to which the EO does not apply. In a sense, the EO can only posture the U.S. government to set a good example.

Ironically, it is DoD that has the greatest measure of influence and control over its particular commercial sector—the Defense Industrial Base. This is one of the reasons that the practices and policies of DoD can more easily exceed the requirements of the EO, and probably the one major reason why the EO exhibits such a light hand on the department.

### *Autonomous Systems and Human-in-the-Loop*

The autonomy of AI-driven machines has long been a political concern, usually captured emotion-laden public debates on "killer robots" able to cause disaster without human controls.[6] The White House's EO reflects the wider public debate on the topic which today is more likely to address police use of robots than the military use.[7] Furthermore, the DoD's often repeated policy of limiting autonomy by keeping a "human-in-the-loop" should in principle serve to discourage military robotic adventurism. Less mentioned, however, are more specific alternatives such as human-on-the-loop (HOTL) supervision, human-adjacent-to-the-loop (HATL) refining, or even human-out-of-the-loop

(HOOL) automaticity, which pose increasingly more distant human direct control of what the AI-trained and controlled machines ultimately do.[8]

It is a sensible assessment of possible future combat to attempt to envision the conditions under which human-in-the-loop intentionally, or not, becomes human-on-the-loop only.[9] That is, when having a human deeply embedded in assessing a situation and then choosing and initiating the next steps becomes a too slow, inaccurate, or risky process. Human-on-the-loop then becomes a human initiating and then monitoring a process by which the AI-controlled machine then assesses, chooses, and then autonomously executes the sequence of steps towards an objective. What then is acceptable when, due to the exigencies of super rapid combat with AI-enabled adversaries, even human-on-the-loop becomes too slow or risky? The next iteration is the human-adjacent-to- or human-near-the-loop (HNTL) where the human is removed from real-time monitoring to episodic monitoring conditions by which interactive machines collectively share information, assess, choose, and initiate steps in operational processes. Finally, a devolution to human-out-of-the-loop or, at least, humans-on/offing-the-automatically-triggerable-autonomous-otherwise-sealed-loop could easily become the only militarily sensible choice against adversaries doing the same.[10] It is not a slippery slope, but possibly an evolution in practice from the machines needing human interaction to their not needing humans or not having many entry points by which humans could interfere. Having the slow and nonexpendable humans interfering in rapid-fire, drone swarm, battle decisions might not be sensible in high-end, high-stakes warfighting in a highly contested electromagnetic warfare-rich environment, other than to centrally shut systems down.

Deputy Secretary of Defense Kathleen Hicks underscores the importance of ensuring that the "appropriate level of human judgment" be "flexible."[11] DoD policy on Lethal Autonomous Weapons Systems (LAWS) does not expressly bar development or use of these capabilities. The key distinctions lie in the definition summarized by the Congressional Research Service; LAWS "are a special class of weapon systems that use sensor suites and computer algorithms to independently identify a target and employ an onboard weapon system to engage and destroy the target without manual human control of the system."[12] To "independently identify" targets and act to harm "without manual human control" defines the LAWS, and therein also lies some of the flexibility in defining "in-the-loop" noted by Dr. Hicks. The human remains "in-the-loop" if the targets cannot be selected without human input, or if control is delegated under human supervision with the option to cancel the mission at key points in its execution.

For example, for the recently announced "Replicator" program, which is designed to produce thousands of surveillance, disrupting, or lethal drones, the human 'in-the-loop' would control overarching instructions to the drone swarm, yet it is unclear what else the human controls in these drone operations.[13] The human-loop evolution is still evolving within DoD along with everyone else; however, the Administration's new EO makes no mention of this challenging spectrum of human involvement in targeting.

### *Russia, The People's Republic of China, and Military AI*

Adversaries and vendors also get a vote in the future of trust with AI and DoD. Their choices could increase environmental pressures on the U.S. military to move faster, to more precisely define what falls within (and without) the outer bounds of acceptable and unacceptable flexibility in what human-in-the-loop means, and to unwittingly expand the operational choices of automation, despite the best of ethical intentions of DoD leaders. There is no reason to assume that aggressive autocratic nations like Russia or China have any incentive to maintain the human-in-the-loop 'principle' concerning the use of deadly force in autonomous systems. Russia's Vladimir Putin has urged international controls on other states while asserting the need for Russia to ensure America has no monopoly on AI.[14] Rapidly adapting commercial AI for use in battle is hard and time-consuming. Commercial developers, absent regulations or other oversight, do not spend the time for careful consideration of the longer-term systemic ethical consequences in democracies, let alone in today's war-driven Russia.[15] And Putin is incentivized to move quickly given his stated position that the nation who "rules AI" will "rule the world."[16]

China's Xi Jinping repeatedly has expressed strong ambition for China to lead the world in artificial intelligence and other technologies perceived as critical to the national socio-technical-economic systems of democratic systems. The Chinese 2017 'New Generation AI Development Plan' projected that China would be the 'major AI innovation hub of the world' by 2030.[17] At the same time, China has a robust history of technological theft through cyber attacks, hostile buys, and bullying economic or political coercion.[18] One after another major cyber attack has been attributed publicly to Chinese state-sponsored hackers across major underlying operating systems.[19] There is no reason to presume any change in China's extraction operations as part of China's relentless centrally fueled AI development program.[20] Nor is there any evidence that the Chinese approach to gaining dominance in or employing artificial intelligence – whether in espionage, economic theft, or combat – will be any different than its admittedly creative cyber extractive pursuit of other technology targets.[21] The one silver lining is the natural conservative bent of the Chinese military to want to be sure a new technology works as planned before it is used extensively.[22] Beyond that functional certainty, however, China's embrace of AI reveals no more ethical constraints than China has shown with cyber in general when opportunities emerge to disrupt democratic targets in or outside of combat.

### *Industrial Base and Dual-Use*

Vendors also get a vote. The principles of the White House AI EO, as the DoD's AI Ethics are voluntary for any commercial entity unless one is already a contractual partner to the DoD in the Defense Industrial Base (DIB).[23] The non-DIB commercial AI industry conducts the vast majority of AI development and production, especially the massive foundation models and their products, but these firms are not obligated to ensure ethical standards through their development or deployment process. Given what is likely to be available within DoD's accelerated time limits, ethically integrating AI effectively inevitably will force choices between

ensuring ethical standards or missing deadlines. Rigorous observance of ethical standards throughout the AI development cycle will be more difficult than assuring the cyber security of the software development process, and the latter is largely not yet anything close to a success.[24] DIB contractors must adhere to DoD ethical principles in order to obtain and retain DoD contracts, but they are by and large not the major developers or vendors of cutting edge AI despite their glossy advertisements making timelines, ethics, and bleeding edge quality of DoD's AI purchases open to question and possible distrust, despite assurances today.[25]

Impacted by the dominance of commercial actors are challenges presented by the dual-use of AI potentially facilitating Chemical, Biological, Radiological & Nuclear (CBRN) development by terrorists or non-experts, regardless of whether humans are in, on, nearby, or out of the loop. Again, although the EO's dual-use provisions will affect private commerce relating to defense, most of the directives only affect other government agencies. DoD and the intelligence community are already deeply invested in following well-known and cyber-enabled CBRN connections, but public sources do not reveal whether this issue will have sufficient AI expertise or dedicated focus.

### *Demands on DoD to Lead AI Policy*

The EO provides little guidance as to these serious external pressures on DoD, rather it has only several demands on the department. The Secretary of Defense is directed to "capitalize on AI's potential to improve United States cyber defenses" (for national security systems) and (along with Department of Homeland Security) "develop plans for...an operational pilot project to identify, test, evaluate, and deploy AI capabilities, such as large language models, to aid in the discovery and remediation of vulnerabilities in critical United States Government software, systems, and networks." The Secretary is also to contract, in consultation with the National Security Advisor and the Office of Science and Technology Policy, with the National Academies of Science, Engineering and Medicine on a study to assess ways in which AI can increase or reduce biosecurity risks. Finally, the Defense Secretary and Chairman, Joint Chiefs of Staff are also members (along with other Cabinet Departments and numerous Presidential advisors) of a White House Artificial Intelligence Council to formulate and implement AI policies.

Also, the White House's National Security Advisor must develop a "memorandum" addressing "the governance of AI used as a component of a national security system or for military and intelligence purposes. The memorandum shall take into account current efforts to govern the development of AI for national security systems."

The bottom line is that the EO's authors appear to accept the reality that DoD already is addressing the EO's key issues in the following: 2022 DoD Responsible AI Implementation and Strategy, 2023 DoD Data, Analytics, and Artificial Intelligence Adoption Strategy, and the various service's AI policies embedded with ethics throughout acquisition/development and deployment/retraining. Other issues can be associated with DoD initiatives such as the

following examples (the list is not exhaustive): promoting innovation and competition (e.g., the "Tradewinds" initiative under DoD's Chief Digital and AI Office); promoting U.S. AI talent ("500 new researchers by 2025" and various DoD AI education efforts); development of training resources (various DoD AI education efforts); establishing new service-connected AI Research Institutes (Department of the Air Force MIT AI Accelerator, Department of the Army's work with Carnegie Mellon University, Department of the Navy efforts to create a DON AI Accelerator homed at NPS); grants for AI Tech Sprint competitions (support for Defense Innovation Unit efforts, Naval Innovation Exchange (NIX), The Navy's Task Force 59, the Naval Applications of Machine Learning (NAML) conferences, and the equivalents across all services).[26] In short, these DoD initiatives help explain the White House's light hand relative to DoD.

The question is whether sufficient thinking has been put into what happens as to the civilian sector spillover when the forcing functions of AI-enabled combat advantages, adversaries' AI advances, and AI vendors' ethical or security neglects inevitably come into play. The impact of this trifecta will be further affected by other technologies changing at warp speed: quantum, nano, and fusion. The AI era is just beginning to outweigh cyber in its transformational effects, including in conflict. Executive orders of the future governing responsible AI must recognize and appreciate the impact that adversarial scenarios will pose for guardrails desperately needed in non-DoD contexts. While the EO did not signal that coming discussion, DoD may need a stronger hand given the bleed-over effect its advances will have on non-DoD capabilities. The military leaders need at least senior leaders' foresight and more collaborative oversight into what are today's exceptionally uncomfortable topics; these dilemmas are likely to be tomorrow's ugly realities.

## DISCLAIMER

The views expressed here are those of the authors alone and do not necessarily represent the views, policies, or positions of the U.S. Government, Department of Defense or its components, to include the Department of the Navy or the U.S. Naval War College.

## NOTES

1. U.S. Department of Defense, *U.S. Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway*, June 2022, https://www.ai.mil/docs/RAI_Strategy_and_Implementation_Pathway_6-21-22.pdf.
2. U.S. Department of Defense, *Remarks by Deputy Secretary of Defense Kathleen H. Hicks on 'The State of AI in the Department of Defense' (As Delivered)*, November 2, 2023, https://www.defense.gov/News/Speeches/Speech/Article/3578046/remarks-by-deputy-secretary-of-defense-kathleen-h-hicks-on-the-state-of-ai-in-t/.
3. U.S. Government Accountability Office (GAO) "Artificial Intelligence: Agencies Have Begun Implementation but Need to Complete Key Requirements" (December 12, 2023) https://www.gao.gov/assets/d24105980.pdf
4. "Executive Order 14110 of October 30, 2023," *Federal Register* 88, no. 210 (November 1, 2023), 75191, https://www.govinfo.gov/content/pkg/FR-2023-11-01/pdf/2023-24283.pdf.
5. Executive Order 14110 of October 30, 2023," 75210.
6. Charli Carpenter, "How scared are people of "killer robots" and why does it matter?" *Open Democracy*, July 4, 2013, https://www.opendemocracy.net/en/how-scared-are-people-of-killer-robots-and-why-does-it-matter/; Jay Stanley, "It's simply too dangerous to arm robots," *ACLU*, December 16, 2022, https://www.aclu.org/news/privacy-technology/its-simply-too-dangerous-to-arm-robots.
7. Stanley, "It's simply too dangerous to arm robots."
8. See discussion in George Galdorisi and Sam J. Tangredi, *Algorithms of Armageddon: The Impact of Artificial Intelligence on Future Wars* (Annapolis, MD: Naval Institute Press, 2024), 125-144.
9. Galdorisi and Tangredi, 74-75;
10. Eric Jensen and Carolyn Sharp, "Reentering the Loop," *Articles of War*, February 9, 2022, https://lieber.westpoint.edu/reentering-the-loop/.
11. Brad Dress, "Pentagon's new AI drone initiative seeks 'game-changing shift' to global defense," *The Hill*, September 6, 2023, https://thehill.com/homenews/4189864-pentagons-new-ai-drone-initiative-seeks-game-changing-shift-to-global-defense/.
12. Congressional Research Service, Defense Primer: *U.S. Policy on Lethal Autonomous Weapons Systems*, February 1, 2024, https://crsreports.congress.gov/product/pdf/IF/IF11150#.
13. Dress, "Pentagon's new AI drone initiative seeks 'game-changing shift' to global defense." For a critical view of 'Replicator,' see Sam J. Tangredi, "Replicate Ordnance, Not Cheap Drones," U.S. Naval *Institute Proceedings* 150, no. 3 (March 2024), https://www.usni.org/magazines/proceedings/2024/march/replicate-ordnance-not-cheap-drones.
14. Samuel Bendett, "Putin Urges AI Limits — But for Thee, Not Me?" *Defense One*, December 3, 2020, https://www.defenseone.com/ideas/2020/12/putin-urges-ai-limits-thee-not-me/170458/; Guy Faulconbridge, "Putin says West cannot have AI monopoly so Russia must up its game, Reuters, November 24, 2023, https://www.reuters.com/technology/putin-approve-new-ai-strategy-calls-boost-supercomputers-2023-11-24/.
15. Ionut Arghire, "Over a Dozen Exploitable Vulnerabilities Found in AI/ML Tools," *Security Week*, November 17, 2023, https://www.securityweek.com/over-a-dozen-exploitable-vulnerabilities-found-in-ai-ml-tools/.
16. James Vincent, "Putin says the nation that leads in AI 'will be the ruler of the world,'" *The Verge*, September 4, 2017, https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world; Galdorisi and Tangredi, 1-4.
17. Ulrich Jochheim, European Parliamentary Research Service, "China's Ambition in Artificial Intelligence," European Parliament (official site), September 2021, https://www.europarl.europa.eu/RegData/etudes/ATAG/2021/696206/EPRS_ATA%282021%29696206_EN.pdf.
18. Kantaro Komiya, "US, Japan authorities warn of China-linked hacking group BlackTech," *Reuters*, September 28, 2023, https://www.reuters.com/technology/us-japan-authorities-warn-china-linked-hacking-group-blacktech-2023-09-28/; Daryna Antoniuk, "China-based spies are hacking East Asian semiconductor companies, report says," *The Record*, October 6, 2023, https://therecord.media/china-budworm-apt27-east-asia-semiconductor-companies.
19. "Fast Thinking: A turning point on Chinese hacking," *Atlantic Council*, July 19, 2021, https://www.atlanticcouncil.org/content-series/fastthinking/fast-thinking-a-turning-point-on-chinese-hacking/.
20. Jane Cai and Willian Zheng, "China launches new data agency as ambitions in AI and digital economy soar," *South China Morning Post*, November 20, 2023, https://www.scmp.com/news/china/politics/article/3242067/china-launches-new-data-agency-ambitions-ai-and-digital-economy-soar.

## NOTES

21. U.S. Cybersecurity and Infrastructure Security Agency, "Cyber Security Advisory: People's Republic of China State-Sponsored Cyber Actor Living off the Land to Evade Detection," May 24, 2023, https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-144a.
22. David Roza, "How China's PLA Could Use Tech Like ChatGPT, And What Could Hold It Back," *Air & Space Forces,* August 24, 2023, https://www.airandspaceforces.com/china-generative-ai-chatgpt/.
23. Congressional Research Service, *Defense Primer: U.S. Defense Industrial Base*, April 17, 2023, https://crsreports.congress.gov/product/pdf/IF/IF10548.
24. Ionut Arghire, "Over a Dozen Exploitable Vulnerabilities Found in AI/ML Tools."
25. U.S. Cybersecurity and Infrastructure Security Agency, "Defense Industrial Base Sector: Council Charters and Membership," n.d. (accessed March 4, 2024), https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors/defense-industrial-base-sector/sector-charters-and-membership.
25. U.S. Cybersecurity and Infrastructure Security Agency, "Defense Industrial Base Sector: Council Charters and Membership," n.d. (accessed March 4, 2024), https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors/defense-industrial-base-sector/sector-charters-and-membership.
26. Joseph Clark, "DoD Releases AI Adoption Strategy," *U.S. Department of Defense*, November 2, 2023, https://www.defense.gov/News/News-Stories/Article/Article/3578219/DoD-releases-ai-adoption-strategy/.