Toward a Balanced Approach: Bridging the Military, Policy, and Technical Communities

Arun Seraphin and Wilson Miles

The *2022 U.S. National Defense Strategy* states that one of its goals is to "modernize the systems that design and build the Joint Force, with a focus on innovation and rapid adjustment to new strategic demands."[1] Consistent with that effort, the U.S. Department of Defense (DoD) has identified artificial intelligence (AI) as a technology with disruptive potential for defense capabilities and highlighted it as a critical technology area for enhanced attention and investment. It is well understood that AI, machine learning (ML), and autonomy, are all poised to drive military technological advantages. The National Security Commission on Artificial Intelligence noted that "the ability of a machine to perceive, evaluate, and act more quickly and accurately than a human represents a competitive advantage in any field—civilian or military."[2] As such, AI holds incredible promise to improve the ability and function of nearly all defense systems and operations.

Given the tactical and strategic value of the technologies and the proliferation of threats, the military continues to explore the development of new autonomous technologies to execute national security missions. Autonomous systems can process and use tactical data and sensing information with speeds impossible for human operators and represent an excellent opportunity to relieve human operators from some "dull, dirty, and dangerous" activities.[3] However, the nature of

Arun Seraphin, Emerging Technologies Institute, Virginia, United States (aseraphin@ndia.org)

Wilson Miles, Emerging Technologies Institute, Virginia, United States (wmiles@ndia.org)

Ethics & International Affairs, 37, no. 3 (2023), pp. 272–286.

doi:10.1017/S0892679423000321

AI and the autonomous defense platforms and weapon systems that AI may enable sparks concerns about their ability to enable operations outside traditional international norms and current or impending formal international agreements, leading to calls for new international agreements to restrict or ban lethal autonomous weapon systems.[4]

Discussions on the promise of autonomous defense systems point to several technical issues that may lead to complex policy and ethical considerations. These include—but are not limited to—how systems deal with compromised, biased, and synthetic datasets; the level and manner in which human interaction with these systems both controls the systems' behaviors and teams with them to execute missions; the speed of machine learning resulting in modifications of system behaviors; the metrics used by the systems to optimize their performance; the predictability and reliability of the systems especially in dealing with rare or unexpected situations; the certification of developmental autonomous systems prior to their use; and the processes used to explain systems behaviors during mission activities.

The development of new technologies that enable autonomous weapon systems poses a challenge to policymakers and technologists trying to balance military requirements with evolving DoD guidance, international obligations, commercial market considerations, and traditional ethical norms. In the military context, this is complicated by standard military technology acquisition processes that begin with threat- or technology-based requirements that drive the development and production of new systems. The process of developing requirements is optimized to meet military operational goals and sometimes operates without close coordination with policymakers or technical experts. It is further complicated by China and Russia actively developing autonomous, AI-enabled weapon systems, since both countries have a track record of not considering the same kinds of ethical issues as those being discussed by the United States.

The global and dual-use nature of this sector and the speed of technological development will further complicate the development of the technologies and policy framework. Commercial and global innovators and customers of AI systems will establish their own preferences and ethical norms outside the needs or concerns of the military. This will drive AI–product development, sales, and proliferation, which in turn will affect the capabilities transferred to military systems.

Nevertheless, we argue that developing a workable and realistic regulatory framework governing the allowable use of lethal autonomous weapons by the military and the artificial intelligence that underpins autonomy is possible. It will

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

require a coordinated effort of the regulatory community, commercial and military technologists (including independent auditors), and military operators to create requirements that reflect the global proliferation and rapidly evolving threat of autonomous weapon systems.

Recognizing the challenges associated with governing the development and use of the technology as well as the need for transparent policies, the DoD has released multiple guidance documents articulating how it views the development and use of automated systems in the age of artificial intelligence (AI).[5] In 2018, the DoD published a summary of its artificial intelligence strategy outlining how it intended to leverage AI in its missions. The document lists five pillars that provide an overview of the DoD's efforts to accelerate AI adoption including, "leading in military ethics and AI safety" on the use of AI in a lawful and ethical manner.[6] Additionally, in 2020, the DoD adopted five ethical principles of AI—suggesting that it should be "responsible, equitable, traceable, reliable, [and] governable"—based on dialogue between policymakers, technologists, and operators.[7] In 2022, the Biden administration signed the DoD's *Responsible Artificial Intelligence Strategy and Implementation Pathway*, which outlines the department's approach to implementing the five DoD AI ethical principles.[8] In 2023, the DoD updated Directive 3000.09, "Autonomy in Weapon Systems," which clarifies the roles and responsibilities of the technical, policy, and military communities that manage autonomous systems' maturation and eventual use.[9]

While the DoD appears to be demonstrating a firm commitment to remaining the global leader in developing and deploying new autonomous systems in a lawful manner, this essay draws attention to key considerations for the technical, policymaking, and military communities when thinking through the opportunities and risks of autonomous weapon systems that will help those commitments be fully realized. First, it sets the stage by using two historical case studies to demonstrate that: (1) the lack of coherent dialogue between the technical and policy communities can result in security, ethical, and legal dilemmas, and (2) bridging the gaps between the military, technical, and policy communities can lead to useful technology with appropriate constraints that balances the needs of all communities. While there are numerous examples illuminating both good and bad coordination, these two case studies were chosen for their clear-cut depiction of how outcomes can be shaped by dialogue.[10]

The essay then builds on the two case studies by categorizing key takeaways from interviews conducted with twelve subject matter experts (SMEs) from the

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

commercial and defense industry as well as academia—many with previous senior leadership experience within the government or military. The themes from the interviews further the conversation by providing concrete considerations that a policymaker or technical director might encounter when deliberating over autonomy development and deployment. The SMEs were selected based on their backgrounds in technical autonomous system development or AI defense policy formulation, including direct experience supporting the formulation of U.S. and international AI and autonomy standards, such as the White House AI Bill of Rights, the IEEE's Standard for Transparency of Autonomous Systems, and the DoD's 5 Principles of Artificial Intelligence Ethics. The SMEs were asked to describe their experience; how constraints could be built into development; and the relationship between the technical, policy, and operational communities.[11]

We then conclude with recommendations that the DoD can pursue to bridge the technical, operational, and policymaking communities. The recommendations are informed by the lessons learned from the case studies and key takeaways from the interviews with SMEs. The goal is to provide the DoD with concrete steps for developing organizational structures or processes that will incentivize engagement across communities.

Historical Case Studies

Case Study 1: Misaligned Policy—Gain-of-Function Research

In October 2014, the White House Office of Science and Technology Policy and the Department of Health and Human Services announced that the U.S. government launched an effort to assess the potential risks and benefits of gain-of-function research (GOFR)—a type of virology that includes studies on methods to affect a pathogen's ability to cause a disease.[12] Simultaneously, the National Institutes of Health instituted a pause in funding for "any new studies involving these experiments."[13] The pause was originally a response to a series of biosafety incidents at federal laboratories, which did not actually involve GOFR but raised concerns about laboratory safety and security as it relates to pathogen research.[14]

During the review, there were two entities responsible for assessing the current state of GOFR. First, the National Science Advisory Board for Biosecurity (NSABB) was tapped as the official federal advisory body for providing advice on oversight of GOFR.[15] Its responsibility was to develop a report with concrete

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

recommendations to inform the development and adoption of new U.S. government policies regarding GOFR. Second, the National Research Council (NRC) of the National Academies of Sciences, Engineering, and Medicine was asked to convene two public symposiums with two goals: (1) assessing the issues associated with GOFR, and (2) evaluating the draft recommendations presented by the NSABB.[16] The public discussions were key to the eventual adoption of new guidelines for the evaluation of GOFR and funding considerations. The funding pause was lifted in 2017 after three years.

During the symposiums, the technical participants reiterated that the term "gain-of-function" was overly broad terminology that unintentionally touched other research projects.[17] A defining issue in understanding GOFR is that among scientists there is no concrete definition of the term, and it has expanded to include several different types of work, many of which may not be considered high risk.[18] For example, the ban's original implementation paused work using GOFR to study vaccine efficacy on transmissible strain mutations of the seasonal flu, despite the policy's intent not to impact work on naturally occurring viruses. One of the NSABB's primary findings was that only a small subset of GOFR entails risks that are significant enough to warrant oversight.[19] As a result, a multiagency scientific review process was established to ensure that GOFR receives appropriate oversight based on technical analysis of risks and benefits.[20] The review also determines whether the research is acceptable for funding and guides subsequent oversight at the federal and institutional levels.

This case study provides several lessons. First, technologists and policymakers need to work together to precisely define technical terms and policy statements. This is especially true given the complex technical jargon and commercial terminology used when discussing AI capabilities. Second, the two symposiums hosted by the NRC demonstrate how public inclusion in federal regulation can lead to beneficial policy changes. Transparency and a public hearing of views help demystify complex technologies and their implications. Lastly, when considering the ethical implications of scientific research, it is difficult to draw a sharp line between ethically "acceptable" and "unacceptable" research. Research in complex fields like GOF and AI will not be clearly delineable as ethical or unethical in many cases. Instead, it will lie somewhere on a spectrum of activities that attempt to measure their consistency with ethical standards. It is important to evaluate all such research based on its potential positive payoff, and to design oversight mechanisms in accordance with risk.

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

Case Study 2: Aligned Policy—F-35 Joint Strike Fighter Exports

The F-35 is the first U.S. stealth fighter to be offered for export and is currently operational in nine different militaries. Exportability is often a key piece of successful defense acquisition efforts, both serving military operational needs in developing coalition capabilities and providing significant cost savings and economic benefits for industry and governments alike. The government and industry team producing the F-35 fighter plane regularly adapts systems to be sold to foreign partners to suit the foreign nation customer's operational needs and to ensure compliance with U.S. export policies, including the Arms Export Control Act of 1976 and the Foreign Assistance Act of 1961. These changes frequently relate to the technical capabilities of weapons or the foreign customer's military operational use concepts, and are also coordinated with policymakers in Congress, Department of State, and the DoD.[21]

From the beginning of the F-35 Joint Strike Fighter program, the DoD and the international community recognized the potential benefits of partnering to decrease costs, share technical knowledge, and provide partners with the opportunity to acquire a fifth-generation strike fighter. Eight countries signed a Memorandum of Understanding (MOU) that laid out the roles, responsibilities, processes, and procedures as well as resource commitments among all signatories.[22] The original countries both collaborated and contributed financially to the development effort.[23] In this MOU, an oversight structure called the JSF Executive Steering Board (JESB) was established to bring together technical experts, operational users, and individuals with policy backgrounds—including international partners—in a single F-35 program office.[24] Each partner country also established its own JSF program office, which has served as an internationally connected coordination structure for the F-35 enterprise. This structure is used to evaluate new technical opportunities and operational challenges, balancing them with policy and even political considerations. There were also advisory groups and functional working groups established, which gives the program a process for incorporating new technologies into the base design.[25]

Program modifications can be hardware or software changes and can appropriately increase or decrease capabilities relative to U.S. aircraft. For example, U.S. international partner F-35 export versions are designed to have a larger radar signature due to Pentagon officials fearing the spread of sensitive stealth technology. Other technological changes include hardware modifications, such as to Norway's F-35 fleet, which uses unique tires and a drag chute to operate effectively in icy

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

conditions.[26] There are also software changes, such as to Israel's test F-35I that is integrated with locally developed electronic warfare and communication systems.[27] Both of these variants are designed to meet the customer's operational needs while still meeting export control regulations.

This cooperation across governments and offices is facilitated by the JESB, which is set up to float issues from the technical community to senior leadership, which includes individuals from all communities and partner countries. There is also a constant dialogue between the operational, technical, and policy communities. Oftentimes, technical decisions are informed by political considerations that are directly fed to the F-35 program office by policy experts. For example, policy experts advise industry and the military on how political climates may shape the availability of certain critical components and on overall technology sourcing and sharing opportunities. A robust and continuous dialogue between these communities allows for the maximum sharing of technologies while enabling partner nations to protect their most important technical and operational secrets.

This techno-policy dialogue enables partner countries to request specific modifications, such as the changes to the F-35s provided to the Norwegian and Israeli air forces. It also allows for technical expert exchanges. The policy community is key to bringing technical foreign partners to the United States to work on operational tests by ensuring that no political issues are overlooked. By bringing in the policymaking community in a coordinated fashion, the F-35 program office is reducing barriers to making needed technical changes to these complex systems, thereby improving operational capabilities for each partner.

This relationship between the technical, user, and policymaking communities directly controls the progression of the overall platform. Ultimately, this is a direct result of how the F-35 program office is organized and the regulatory framework it uses, leading to the most advanced stealth fighter operating from twenty-seven bases worldwide and being used by nine countries. While selling such an advanced fighter to foreign countries requires untangling a thicket of technical, policy, and operational concerns, the exportability of F-35s and the program's coalition partnership demonstrates that bridging the military, technical, and policy communities can lead to outcomes that balance the needs of all communities while keeping the operational user in mind. Well-informed policies originate from early and continuous engagement, which positively influence a range of activities throughout the technology's research, development, testing, manufacturing, and operational use life cycle.

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

Expert Interviews: Technology and Policy Considerations for AI-Enabled Autonomous Systems

Interviews that we conducted with twelve AI and autonomy experts highlight some key opportunities, guidelines, and concerns of policymaking and technology communities on autonomous system development and deployment that can support efforts to align technical realities, policy considerations, and user needs. Interviewees noted that this alignment creates an environment that allows for constructive and continuous feedback between the sectors. The interviewees from the technical community also expressed an eagerness for the policy and user community to have a deeper understanding of the technical components of AI. Similarly, there is a lack of education on policy, dual-use technologies, regulatory, or ethical considerations in the technical training and education of most scientists and engineers. There is also insufficient representation of each of these groups in the typical discussion forums. Oftentimes, the technical inputs are overwhelmed by those from policy and the military. Moreover, there are few opportunities for sustained and detailed dialogue and collaboration with equal input from all. Experts from the interviews did not, in principle, have a negative view of policy constraints on the development and deployment of AI-enabled systems. Instead, they recognized the need for a policy governance structure to support the industry's ability to create innovative capabilities and uses, and to provide new AI-enabled systems to commercial and defense customers worldwide. The complexity of responsible autonomy development and deployment illustrates the importance of early and continuous engagement across communities.

The following takeaways provide insight into what the technical and policymaking communities consider fundamental to the progression of responsible autonomous development. These key takeaways are: (1) data governance is highly important for aligning a system's goals and actions with human intent, (2) risk tolerance for systems is understood differently for different missions, and (3) diverging definitions or understandings of what constitutes "autonomy" may impede the possibility of regulation or be incongruent with current regulation. Each category validates the need for robust communication between each community, given the challenges not just to the developers and operators but to the policymakers as well.

Data Governance

Governance of the datasets that AI-enabled systems use and learn from is fundamental to responsible development. Understanding the biases and sources of the

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

data is critical to ensuring that its use by AI-enabled systems is consistent with ethical norms and policies. In edge cases, these biases from supervised learning can lead to unforeseen errors, such as target misidentification. Governance activities need to extend to regulation of the content and use of synthetic data sets and system modeling tools, which are increasingly used as a part of generative AI efforts to train systems in situations where real-world data is limited or unavailable.

The policy community also needs to be involved in the development of metrics by which AI-systems optimize their performance based on their programmed goals. These metrics control how AI-enabled systems use data to support active learning systems attempting to optimize their performance. If these metrics are not aligned with and constrained by policy considerations, the systems will move into potentially unintended and unregulated behaviors.

Multiple interview subjects pointed out that the technology sector is developing its own efforts at ethical design and use of AI-enabled systems at the company level, in standards development, and through technical associations. Clear and well-informed policies lead to better technology, more predictable markets, and an ability to expand into larger markets. The Institute of Electrical and Electronics Engineers (IEEE) Global Initiative on Ethics of Autonomous and Intelligent Systems provides an example of the technical community constructively engaging in a discussion of AI development and ethical principles. The IEEE's *Ethically Aligned Design* report is a tangible product informed by IEEE's global community, which seeks to provide stakeholders involved in the design and deployment of AI-enabled systems training to enhance considerations of ethical issues during technology development.[28]

Mission-Level Risk Tolerance

Interview subjects from both industry and government backgrounds noted that policy discussions rarely form around realistic mission-use cases or are rooted in specific concepts of operation. Much of the current discourse surrounding autonomous weapon systems occurs at an abstract level, where debate tends to focus on theoretical examples. They recommended that mission-based use cases should be used to help shape technology development as well as the policy questions that arise from deployment.

AI can be a tool to support decision-making, as it can provide more accurate assessments needed for target discrimination or identification of friend and foe.

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

AI capabilities increase the ability of decision-makers to make precise judgments in a conflict that align more with policy considerations than they do with either unguided munitions or stressed military personnel with limited training. Trust is critical to effective human-machine teaming, and it is likely that no other profession or institution relies more on trust than the military. Therefore, AI's use should be weighed among other available options.

At the mission level, authorization of operational uses of autonomous systems will often depend on risk tolerance. Data collected from mission simulations can inform robust verification and validation and lead to quantifiable performance levels within the tested scenarios that all parties can assess for acceptable risk. The complexity of autonomous system behaviors will limit the ability of technology developers to give complete assurance of the predictability of system behaviors in extreme situations and edge cases.

Lack of Agreed-Upon Language

Technical experts from academia and industry highlighted that there is no agreement on what constitutes autonomy in a weapons system, or what part of the "kill chain" that connects a sensor to a control system to a weapon should be considered, as regulations are developed on the use of autonomy in defense missions. Furthermore, they noted that there is no clear technically defined boundary between autonomous and semiautonomous systems, which complicates the development of discrete and distinct regulatory frameworks. Once the system is deployed, there is no method for inspection to ensure compliance with existing policies.

The integration of autonomy and machine learning into defense systems is a gradual and continuous process. It is occurring constantly at the software and subsystem level through upgrades and incorporation of incremental technical improvements. This type of gradual improvement is commonplace in the commercial sector, and the military sector will follow the same path. Thus, regulators should not expect to be presented with a new "autonomous system" to evaluate; rather, they will have to deal with semiautonomous systems and a mixture of autonomous and supervised activities for all defense systems over time.

The interviewees pointed out that autonomy will be gradually adopted into all systems as part of continuous and natural technological upgrades of software and subsystems. As such, the current term "AI-enabled system" becomes insufficient, as more machine functions become autonomous and the manifestation of autonomy becomes increasingly nuanced, which makes regulation much more difficult. It is

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

not likely that newly developed systems will be presented as a complete "AI-based autonomous system" to policymakers. Rather, they will need to be continuously assessed and regulated as their capabilities are actively upgraded over time.

Recommendations

The U.S. government and defense industrial base must play an active role in shaping domestic and international development of commercial standards for the use of AI since they will also drive technological development by companies supplying the DoD. International considerations of ethical issues on usage in defense or commercial applications will strongly drive global technical developments that will in turn affect the DoD's access to commercial and global technology. Both the case studies and interviews were used to inform the exploration of the relationship between policy and the development of autonomous technologies; the impact of coherent dialogue between the technical, user, and policy communities on creating or avoiding security, ethical, and legal tensions; and the key technical issues that these communities should address going forward. To continue developing best practices and better integrate the policy and technical communities, the following recommendations should be reviewed by DoD policymakers:

1. Initiate a set of activities convening defense and commercial industry representatives, military operators, academics, and policymakers to study and to promote dialogue on the implications of AI developments in autonomous weapon systems and inform government decision-making or regulatory processes.[29] These activities should include symposia; research; intersectoral personnel exchanges; and training for personnel involved in writing of military requirements, acquisition of systems, and technology development. All of which would be intended to strengthen the workforce and be responsive to the needs of actual ongoing programs and activities.
2. Develop a research agenda and funding plan for technologies and capabilities that improve the development of regulatory controls and still meet needed military operational capabilities, including explainable AI and machine learning; generative AI to explore scenarios that are not easily predictable and might prove problematic for autonomous systems trying to operate within regulatory regimes; faster systems for operators to assert human-in-the-loop control when needed over autonomous systems; assessment of bias in datasets and prediction of potential problems

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

with systems that learn from the datasets; and testing capabilities and protocols to continuously evaluate learning systems to ensure they are operating within policy guidance.

3. DoD should establish cross-functional teams from the operational, acquisition, policy and political, and technology development communities focused on the development and use of autonomous weapon systems. This should include putting in place policies that are relevant for different operational tempos and the spectrum of military engagements; for example, the governance for training and exercises may need to be different than that for peacekeeping and humanitarian assistance, as opposed to acceptable use of the technologies during high-intensity combat operations. Pentagon leadership should use the recommendations from these teams to continuously review policies, requirements, technology priorities, and resource allocation.
4. The U.S. government should engage peer competitors, global partners, and the commercial sector in discussions on the ethical use of AI-enabled weapon systems. An international framework informed by the widest set of global players from a range of sectors and perspectives will be key to enforcing any agreed-upon standards and norms on the use of autonomous defense systems.

Given the promise of AI technology, the DoD is correctly seeking to develop and leverage it to support national security missions. However, governing the use of autonomous defense systems will be complex because of the globalized and dual-use nature of the rapidly changing technology, and the constant development and emerging use of these systems by near-peer competitors and threat actors. As has been demonstrated many times in the past, and as was illustrated by the case studies above, regulatory frameworks for new technologies are best developed through comprehensive dialogue between the technical, policy, and user communities. In this promising and accelerating technology area, this engagement is necessary for ensuring regulators, warfighters, and technologists sufficiently balance ethical and policy concerns against technical realities and military requirements.

Notes

[1] U.S. Department of Defense, *2022 National Defense Strategy of the United States of America: including the 2022 Nuclear Posture Review and the 2022 Missile Defense Review* (Washington, D.C.: Department of Defense, October 2022), p. 2, media.defense.gov/2022/Oct/27/2003103845/-1/-1/1/2022-NATIONAL-DEFENSE-STRATEGY-NPR-MDR.PDF.

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

[2] National Security Commission on Artificial Intelligence, *Final Report: National Security Commission on Artificial Intelligence* (Washington, D.C.: National Security Commission on Artificial Intelligence, 2021), p. 7, UNT Digital Library, digital.library.unt.edu/ark:/67531/metadc1851188/m2/1/high_res_d/Full-Report-Digital-1.pdf.

[3] Military examples in this article: www.army.mil/article/154248/dull_dirty_dangerous_mission_send_in_the_robot_vehicle and www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/.

[4] "Lethal Autonomous Weapons Systems (LAWS)," Office for Disarmament Affairs, United Nations, www.un.org/disarmament/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/.

[5] This is in addition to other U.S. government initiatives, including from Congress, the White House, and agencies such as the National Institute of Standards and Technology, which are also developing regulatory frameworks for AI. All of these guidance documents will affect industry and thereby the technologies available to the DoD.

[6] U.S. Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (Arlington, Va.: U.S. Department of Defense, June 2018), p. 5, media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

[7] C. Todd Lopez, "DOD Adopts 5 Principles of Artificial Intelligence Ethics," DoD News, U.S. Department of Defense, February 25, 2020, www.defense.gov/News/News-Stories/Article/Article/2094085/dod-adopts-5-principles-of-artificial-intelligence-ethics/.

[8] DoD Responsible AI Working Council, U.S. Department of Defense, *U.S. Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway* (Washington, D.C.: U.S. Department of Defense, June 2022), media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF.

[9] Office of the Under Secretary of Defense for Policy, DoD Directive 3000.09, January 25, 2023, www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

[10] There are many examples of the impacts of coordination between the policy and technical communities that cannot be detailed in this essay. They include the development of social media, the X-47B unmanned combat air system carrier, and the formulation of the New START Treaty.

[11] Technical interview questions included:

- Are there limits to the kinds of constraints the coder can write? If given the proper time and resources, can programmers build in norms/ethical considerations into the technology?
- What, if anything, are we losing because of ethical/policy constraints (costs, timeline, performance)?
- Learning systems will evolve; how do you create constraints for those types of systems?

Policy interview questions included:

- How do you get your information on laws, regulations, and treaties? How do they shape your autonomous weapons development?
- How do laws, regulations, and international treaties influence your autonomous system development?
- Are there success stories of policy positively influencing the development of a technology?
- What is the best mechanism by which the policy, technology, and warfighting communities can share their expertise to develop the most appropriate policy frameworks?
- How do defense primes and weapon systems developers approach balancing international partner needs with technological feasibility and exportability?
- Can you think of one good story and/or bad story of the policy and technology communities working together (or not working together) to improve an outcome or the regulation of the use of a technology?

[12] "Doing Diligence to Assess the Risks and Benefits of Life Sciences Gain-of-Function Research," President Barack Obama, White House, October 17, 2014, obamawhitehouse.archives.gov/blog/2014/10/17/doing-diligence-assess-risks-and-benefits-life-sciences-gain-function-research.

[13] "Statement on Funding Pause on Certain Types of Gain-of-Function Research," NIH Director, National Institutes of Health, October 16, 2014, www.nih.gov/about-nih/who-we-are/nih-director/statements/

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

statement-funding-pause-certain-types-gain-function-research; and Department of Health and Human Services, National Institutes of Health, *U.S. Government Gain-of-Function Deliberative Process and Research Funding Pause on Selected Gain-of-Function Research Involving Influenza, MERS, and SARS Viruses*, (Washington, D.C.: Department of Health and Human Services) October 17, 2014, www.phe.gov/s3/dualuse/documents/gain-of-function.pdf.

[14] National Science Advisory Board for Biosecurity, *Recommendations for the Evaluation and Oversight of Proposed Gain-of-Function Research*, (Bethesda, Md.: National Institute of Health) May 2016, osp.od.nih.gov/wp-content/uploads/2016/06/NSABB_Final_Report_Recommendations_Evaluation_Oversight_Proposed_Gain_of_Function_Research.pdf.

[15] "Doing Diligence to Assess the Risks and Benefits of Life Sciences Gain-of-Function Research."

[16] Ibid.

[17] Frances Sharples, Jo Husbands, Anne-Marie Mazza, Audrey Thevenon, and India Hook-Barnard, "Gain-of-Function Research: Background and Alternatives," ch. 3 in *Potential Risks and Benefits of Gain-of-Function Research: Summary of a Workshop* (Washington, D.C.: National Academies Press, April 13, 2015), www.ncbi.nlm.nih.gov/books/NBK285579/.

[18] Ibid.

[19] United States, National Science Advisory Board for Biosecurity, *Recommendations for the Evaluation and Oversight of Proposed Gain-of-Function Research.*

[20] Ibid.

[21] Defense Security Cooperation Agency, *Foreign Customer Guide* (Washington, D.C.: Department of Defense, July 15, 2018).

[22] Jerry McGinn and Michael T. Roche, *A "Build Allied" Approach to Increase Industrial Base Capacity* (Fairfax, Va.: Greg and Camille Baroni Center for Government Contracting, George Mason University, June 22, 2023).

[23] The original eight allied countries include the United States, United Kingdom, Canada, Denmark, The Netherlands, Norway, Italy, Turkey, and Australia. In July 2019, the DoD removed Turkey from the development program.

[24] Interview with Vice Admiral Mathias W. "Mat" Winter, former director, Joint Strike Fighter Program, Office of the Secretary of Defense, conducted by authors (July 25, 2023 via Microsoft Teams).

[25] Ibid.

[26] Eric Hehs, "F-35 Lightning Drag Chute," *Code One* 31, no. 1 (August 13, 2014), Lockheed Martin, www.codeonemagazine.com/article.html?item_id=138.

[27] Thomas Newdick, "Israel's Specially-Built F-35I Test Jet Just Touched Down In-Country," Drive, updated November 11, 2020, www.thedrive.com/the-war-zone/37579/israels-specially-built-f-35i-test-jet-just-touched-down-in-country.

[28] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems*, 1st ed. (Institute of Electrical and Electronics Engineers, 2019), standards.ieee.org/content/ieee-standards/en/industry-connections/ec/ autonomous-systems.html.

[29] This might be modeled on activities under the National Nanotechnology Initiative, which has supported the development of safety standards and public educational materials, or on the structure of the F-35 program office.

Abstract: The development of new technologies that enable autonomous weapon systems poses a challenge to policymakers and technologists trying to balance military requirements with international obligations and ethical norms. Some have called for new international agreements to restrict or ban lethal autonomous weapon systems. Given the tactical and strategic value of the technologies and the proliferation of threats, the military continues to explore the development of new autonomous technologies to execute national security missions. The rapid global diffusion and dual-use nature of autonomous systems necessitate a proactive approach and a shared understanding of the technical realities, threats, military relevance, and strategic implications of these technologies from these communities. Ultimately, developing AI-enabled defense systems that adhere to global norms and relevant treaty obligations, leverage emerging technologies, and provide operational advantages is possible. The development of a workable and realistic regulatory framework governing the use of lethal autonomous weapons and the artificial intelligence that underpins autonomy will be best supported through a coordinated effort of the regulatory community, technologists, and military to create requirements that reflect the global proliferation and rapidly evolving threat of

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press

autonomous weapon systems. This essay seeks to demonstrate that: (1) the lack of coherent dialogue between the technical and policy communities can create security, ethical, and legal dilemmas; and (2) bridging the military, technical, and policy communities can lead to technology with constraints that balance the needs of military, technical, and policy communities. It uses case studies to show why mechanisms are needed to enable early and continuous engagement across the technical, policymaking, and operational communities. The essay then uses twelve interviews with AI and autonomy experts, which provide insight into what the technical and policymaking communities consider fundamental to the progression of responsible autonomous development. It also recommends practical steps for connecting the relevant stakeholders. The goal is to provide the Department of Defense with concrete steps for building organizational structures or processes that create incentives for engagement across communities.

Keywords: AI, artificial intelligence, autonomy, autonomous weapons, Department of Defense, policy, machine learning, F-35, gain-of-function research, military ethics

https://doi.org/10.1017/S0892679423000321 Published online by Cambridge University Press