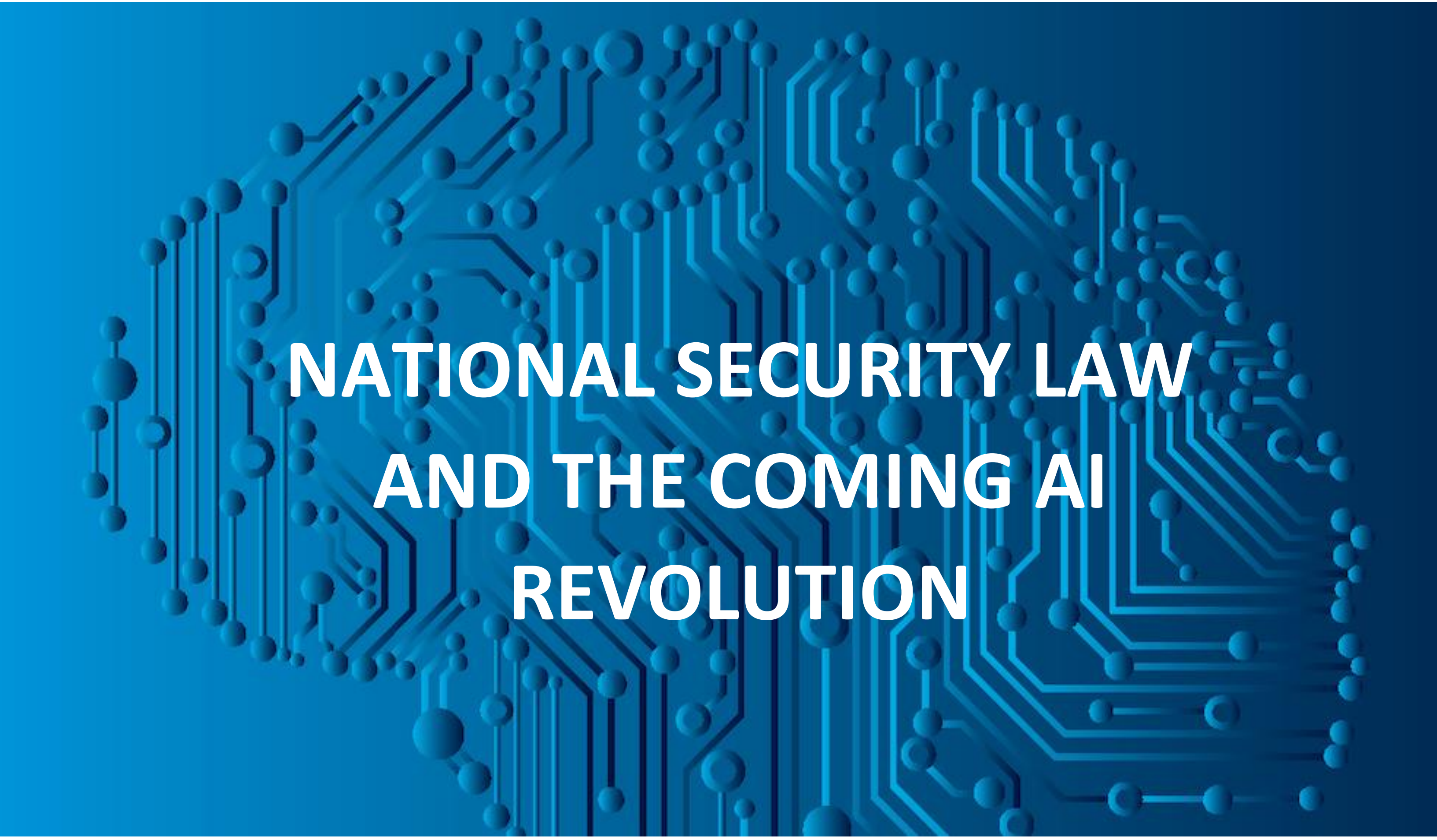


NATIONAL SECURITY LAW AND THE COMING AI REVOLUTION

OBSERVATIONS FROM A SYMPOSIUM HOSTED BY

SYRACUSE UNIVERSITY INSTITUTE FOR SECURITY POLICY AND LAW

AND

GEORGETOWN CENTER FOR SECURITY AND EMERGING TECHNOLOGY

OCT. 29, 2020

Syracuse University
Institute for Security Policy & Law

CSET CENTER *for* SECURITY *and* EMERGING TECHNOLOGY

Contents

Introduction: Present at the Creation

The National Security Commission on Artificial Intelligence (NSCAI) has said, "the development of AI will shape the future of power."[i] The leading academic study of AI and national security concluded in 2017 that "AI has the potential to be as transformative a national security technology, on par with nuclear weapons, aircraft, computers, and biotechnology."[ii] Not to be outdone by the United States, China's State Security Council in July of 2017 committed to spending $150 billion in some manner in the next decade to become the world's leader in AI by 2020.[iii]

The AI wave is breaking over us. We know this now. AI will transform national security practice. In areas like logistics, health management, and intelligence it already has. AI presents distinct opportunities and potential risks. As anyone who has ever shopped, driven a car, or listened to music already knows, AI is here.

AI also brings distinct legal, ethics, and process challenges and risks to the national security space. At the same time, it is axiomatic that case law and statutory law do not keep pace with Moore's law. AI is no exception. What then should the national security legal community do to respond to AI and to prepare for the revolution that will come in: intelligence, decision-making, logistics, and yes, weapons.

The Georgetown University Center for Security and Emerging Technology and the Syracuse University Institute for Security Policy and Law thought that one place to start was with a symposium tailored to national security law practitioners. We hoped to introduce a wider audience of national security generalists to some of the key concepts and issues presented by AI. We also wanted to introduce that audience to some of the leading practitioners and thinkers in the field. Indeed, one of the Symposium panelists is on the NSCAI and another of the panelists is the co-author of the benchmark Belfer/IARPA Study referenced above.

Our goal? To engage the help of a larger audience in crafting, molding, and informing the legal and ethical regime that is beginning to take shape, or in the view of some of our panelists, already exists, to regulate national security uses of AI. To use Secretary of State Dean Acheson's "modest" description of his role in creating the post-World War II international system, today's national security law practitioners have the opportunity to be ‘present at the creation’ – the creation of the legal and ethical regime that will, could, or should govern the use of AI for national security purposes for decades to come. Unless one believes that Google or the Department of Defense should alone make national security policy when it comes to AI, this is a role not just for policymakers and technologists, but also national security lawyers and policymakers, emphasis on the word *national*.

What follows are some of the key observations made at the Symposium by the panelists. We use the term “observations” because we did not ask the panelists to reach conclusions nor find consensus in their discussions or in the presentation of this report. This report, therefore, is a compendium of ideas and thoughts derived from the Symposium for legal practitioners to consider as they proceed to apply general principles and statements of law to specific AI applications and uses. That also means that each of the panelists has plausible deniability. They are free to agree or disagree, associate, or disassociate, with anything found in this report, including any errors. We should also note, as the participants did, that the views expressed were their own and did not necessarily reflect the views of the agencies or entities with which they work or are affiliated.

Roadmap of this Report

The symposium commenced with a presentation on what AI is and how it works to make the technology behind AI accessible to national security generalists. For readers who did not attend the Symposium we collect at the outset of this report some of the general observations made about the constellation of technologies referred to as AI. We then present the key points and observations from each of three panels – AI and the Law of Armed Conflict; AI and National Security: Ethics, Bias, and Principles; and AI and National Security Decision-Making. The Report concludes with a discussion about the role of lawyers, policy-law-technology teaming, and importance of making purposeful ethical and legal choices, which will embed our values in AI applications but also result in more accurate and effective national security tools.

AI: A Constellation of Technologies

> *"AI is not a single piece of hardware or software, but rather a constellation of technologies that give a computer system the ability to solve problems and perform human tasks that would otherwise require human intelligence."*
>
> *NSCAI*[iv]

AI is unlike any technology that has preceded it and its impact, too, will diverge from past innovation patterns. Unfortunately, policymakers often see AI through the lens of fiction. As one panelist noted, the term instantly brings to mind "the fantastic, the terminator, the killer robot," distracting from real world, complex issues including AI bias, verification, and explainability. The gulf between perception and reality only slows our progress in developing sound AI policy and law. Throughout the symposium, panelists underscored the need to focus policy and law on actual applications of AI and those issues that can be pragmatically solved today.

Panelists sought to establish 'what is AI,' what delimits this policy discussion, and what technologies create artificially intelligent systems. Throughout the symposium panelists noted that:

- "Intelligence is the art of prediction. All intelligent activities, whether biological or electronic, involve immensely fast analysis and prediction based on learned experience and data." Artificial intelligence is no different.
- There is no 'one definition of AI.' Several panelists noted that there are boundless definitions and that for policy to progress, our focus should be placed not on a clear-cut definition but on all the technologies we collectively call AI today. Concurring with the NSCAI's definition noted at the start of this section, AI represents a "constellation of technologies." It is "like electricity;" an "enabling technology" that is broadly applicable and can fuel boundless, disparate applications. These include robotics, facial recognition, and image generation. AI cannot be easily pigeonholed.
- Generally, however, AI algorithms are used to "perform human tasks that would otherwise require human intelligence."
- Fundamentally, AI is a computer algorithm designed to "predict optimal future results based on past experience and recognized patterns." This is computer code, which produces predictions, not truths.
- Some panelists concurred that "AI serves as a prediction generator, not a decision generator." Fundamentally, AI is designed to make predictions about what actions should be

taken, and it is the task of policymakers to determine whether that AI or a human has the authority to act on those predictions and make decisions.

Panelists also stepped beyond the technical and definitional adding that:

- "AI also is an ideology" which can be informed by a desire in the minds of some to replace humans and human judgement. Inherent in AI is some form of ideological position about what kind of power can be delegated to a machine-driven process.
- Conversely, this AI ideology may seek not to "replace human intelligence, but to augment it." This can produce results "that [are] better both for humans, for the machines, and for the human machine team."
- AI is also "a goal." Often the value proposition of AI is presented as an idealistic interpretation of not what it is but what it could be. This interpretation includes the potential that AI can be less biased than humans and offer greater than human analytic accuracy. It was also noted that humans are prone to the fallacy of seeing AI as "something that happens tomorrow." Once introduced, once 'magical' AI technologies quickly grow commonplace, raising that bar for what we deem "intelligent."

Frequently, panelists shed light on the real-world national security applications of AI that make up the "AI constellation" including:

- A range of 'boring' application areas such as "logistics, simulation and training." The panelists frequently noted these applications to clarify that AI is often about streamlining the many processes that support intelligence and defense.
- AI is also often used "to assist our operators and military decision makers to make better decisions." These applications often are "narrowly defined" and include task specific AI such as "object recognition."
- Finally, "Intelligent autonomous systems" also represent a "large portion of what most militaries in the world are looking at." The advantages of these systems over human operators include "the machine's ability to adapt in dynamic unstructured and uncertain environments," the ability to adapt at "machine speed," and the "ability to adapt in the presence of overwhelming data input."

Panel 1: AI and the Law of Armed Conflict

The confluence of artificial intelligence, autonomy, and international law is wrought with confusion, making the debate about trends involving technology in weapon systems, and their impact on the laws of armed conflict (LOAC), particularly challenging. As Paul Scharre, author of *Army of None*, noted, "[e]ven setting aside the notion of weapons for a moment, the term 'autonomous robot' conjures up wildly different images, ranging from a household Roomba to the sci-fi Terminator."[v] Despite these challenges, governments have a keen interest in advanced technologies and their current and future impacts on combat operations. Artificial intelligence in weapon systems has the potential to enable combat forces to better understand the environment, to make efficient decisions based on large data sets, to act where appropriate at machine speed, and to act with greater independence from humans in executing the mission in an effective and increasingly humane manner. The *Artificial Intelligence and LOAC* panel presented a practical approach to the legal, policy, and technological debate surrounding technologically advanced weapon systems and their employment on the battlefield.

Key Points and Observations

Some themes that emerged from the panelists' discussion of AI and LOAC were:

- Real World AI Applications
- Weapons Review: Compliance with and the Adequacy of Existing Law
- Ethics and International Competition
- AI and Urban Warfare as a Case Study in Ethics and LOAC compliance
- Operator and Commander Accountability
- Challenges of AI Explainability

Real World AI Applications

Panelists spoke to the importance of framing debate in terms of real-world AI applications, rather than the theoretical:

- The debate surrounding the use of AI has been plagued by unrealistic expectations and misunderstandings. The fear of terminator-like autonomous weapons systems is misinformed. There is little indication that states are interested in completely autonomous AI systems. On the contrary, states are putting considerable effort into controlling every aspect of combat – and therefore AI – on the battlefield.

- AI is already utilized in the military and on the battlefield in less controversial applications. AI is utilized in logistics, simulation, training, and other non-combat areas. On the battlefield, AI is used in narrowly defined areas such as object recognition in image analysis and targeting, to better inform operators and decision makers.
- The ultimate goal of using AI in conflict is to enable operators to make better decisions and support the most efficient military operations.
- The US military aims to augment human intelligence, not replace it. Human-machine teaming remains a main theme in military investments.
- Though AI presents many challenges it is essential that the U.S. continue to research, develop, and ultimately field AI systems. Technological innovations have presented themselves throughout history and have made material changes to the way warfare is conducted. Humans have previously learned to embrace revolutionary technologies such as the cross bow, gun powder, and aviation. AI is no different. It offers tremendous warfighting advantage because it has or has the potential to have the ability to:
    - Adapt in super dynamic, unstructured situations,
    - Adapt at machine speed,
    - Adapt in the presence of overwhelming incoming data, and
    - Function without fear or fatigue.

The Weapons Review Process: Compliance with and the Adequacy of Existing Law

Panelists were asked about the current processes that exist to review and regulate weapons for legality and safety and whether such systems are adequate for AI:

- One panelist was asked to comment on the weapons review process and specifically the implementation of Department of Defense Directive 3000.09.[vi] That Directive establishes DOD policy and assigns responsibility for the development and use of autonomous and semi-autonomous functions in weapons systems, as well as guidelines to minimize failures in those systems. In short, it addresses how autonomous weapons systems are developed and fielded by DOD.
- While Department of Defense Directive 3000.09 has been in place for twelve years, no system has completely gone through the entire review process. With the “expectation that there are systems on the cusp of going through the process,” DOD has been working on how to operationalize the policy and has come to several lessons learned.
- First, it is often unproductive to discuss systems “in the abstract.” Rather, conversations must be grounded in actual exemplar technologies used in exemplar situations. By examining the specifics of the technology and approaches of these systems, analysts can uncover issues that would be otherwise missed but that would create problems later.
- Second, while there is much work to be done in creating an ethics, policy, and possibly legal framework for AI and autonomous systems, we are not missing any laws or gaps in the legal framework “at the highest level.” LOAC concerns how states apply force in warfare but its principles are agnostic as to how states implement force.

- The DOD *Law of War Manual*[vii], however, states that the obligation to ensure force complies with the laws of war is the responsibility of humans and can never be transferred to humans. The DOD has taken this principal as a starting point.
- Below this high level, however, significant gaps – ethics, policy, and legal – exist in *applying* LOAC, or in answering the questions: How do we develop, field, and operate human machine teams where AI-enabled machines inform human decisions as well as make some decisions? How can we employ humans and machines together in a way that is consistent with the law of war?
    - For example, while the DOD has adopted AI ethical principles[viii], those are still too general to adequately regulate autonomous weapons. Policymakers should consider even more specific subsets of ethical principles for both autonomous systems and lethal autonomous weapons systems to ensure that issues particular to those systems are not overlooked by legal and ethical frameworks.
    - Another question, at the crux of 3000.09, is: what is the nature and extent of decision-making that machines are permitted? This is not an entirely new question: many machines already make decisions today as they do things like look for mines underwater or perform surveillance and reconnaissance.
- When asked whether we needed more clear and enforceable treaty obligations to guard against the risk posed by autonomous weapons systems, at least one panelist said no.
    - LOAC's principles for the conduct of hostilities work well, and it is wise that they are technology-neutral because technology is constantly developing. If, however, humanity was to take on the challenge of adopting law specific to autonomous weapons or lethal AI, it would have to do so very carefully to ensure that those laws adequately covered present and future technology. It would be better to figure out how to use AI consistent with the existing law of armed conflict.
    - Moreover, the panelist suggested there is reason to believe our adversaries will try to manage and use technology "in a way that complies with the legal framework we have already."

Ethics and Competition

The panelists discussed whether adopting legal and ethical standards could erode the United States competitive position:

- One panelist pointed out that this is a fallacy. It is not an either-or choice between competitiveness and ethics. Rather, whatever we do to further ethics helps ensure our competitive advantage on the battlefield and in national security generally by improving accuracy and efficacy. The panelist suggested that the United States adopt a framework that asks, "How can we employ humans and machines together in warfare in a way that is consistent with the law of war?"
- Another panelist stated that the United States' adversaries should be thought of as pragmatic. Even if some may care less about civilian harm, they will not use systems so vulnerable that they fire on their own people. Doing so would not be rational.

- Further, rational actors will develop generally effective systems, so even if the United States' view of ethics differs from the views of particular states, that should not close the door on conversations and collaborations with those states. Rather, agreement might be built on shared pragmatic concerns that still prove helpful in protecting people during war.

AI and Urban Warfare as a Case Study on Ethics and LOAC compliance

Urban warfare presents a case where AI might increase our chances of succeeding at an objective while complying with the laws of war. AI technologies such as rapid information processing and computer vision can work to minimize civilian harm and collateral damage.

- Urban warfare is disproportionately destructive to civilians and civilian structures. The nature of a city makes it difficult to distinguish between legitimate and illegitimate targets. Civilian infrastructure, such as roads, bridges, railroads, electricity, and water supplies, overlap with military infrastructure.
- With further research and development, AI may serve as a tool to reduce the civilian harm and minimize the collateral damage, both generally and in the case of urban warfare. Technologies, such as computer vision and rapid information processing, might be leveraged to improve situational awareness for military operators and help distinguish between civilians and legitimate military targets. The same technologies might serve as a tool to assess battle damage. For example, AI can understand and anticipate daily movement within an urban area and assist in the determination of not just where to strike but when, to avoid undue civilian harm.
- AI is unlikely, however, to solve every problem or even most problems that exist in urban warfare settings. History shows that in an urban setting, the situation on the ground can change drastically, and quickly. In the case of a large event, such as the collapse of a bridge or a building, the gathered intelligence informing your operation becomes irrelevant.
- In a rapidly changing environment, the AI systems we have today are not yet safe, secure, or reliable enough to process real-time data, then update themselves and learn in real time, and thus be used for targeting or other immediate decisional support. This is especially true because the enemy will be targeting the AI systems.

Operator and Commander Accountability and AI Explainability

The panel concluded by turning to issues of accountability for and explainability of AI. The panelists sought to shed light on who, if anyone, is accountable for autonomous systems and how operators and commanders can learn to trust and rely on their systems.

Accountability

- The panelists first established that an autonomous system will never itself be responsible or accountable for failures. AI is an enabler, it is electricity, not something that can be put on trial.

    - Ultimately the person responsible for casualties, disproportionate force, or other violations of the LOAC will be the commander that authorized an autonomous system's use.
- To that end, commanders must do everything feasible to ensure these systems are compliant with the laws of armed conflict, especially in ensuring that civilian casualties are not excessive.
- Other humans who might be held responsible might include civilian contractors under product liability and related theories.

AI Explainability

The panel also addressed how a system can be trusted and relied on by commanders to follow LOAC and various principles if its decision-making process is hidden in a 'black box:'

- One panelist suggested that many argue that AI must be explainable to be legal. LOAC, however, does not speak directly to AI explainability. The panelist suggested that complying with targeting principles, such as proportionately and distinction, will probably require that operators have a good understanding of what a system will do, what it is capable of, and what courses of action it might take in a particular operation. But to say operators will have to have a good knowledge of a system creates a significant challenge when we are dealing with machine learning or neural networks that by definition are not easily comprehendible to the operator.
- It is therefore difficult to argue that explainable AI is a legal requirement, though it is ethically and politically important. Certainly, the more we understand AI systems, the better the chances that we will reach consensus to field them.
- There seemed to be some consensus on the panel that the essential question is how the machine functions: whether it operates in the way the operator intends. Explainability, in many cases, serves comfort and trust, which might be achieved in other ways.
    - Some AI applications demand explainability, such applications where the law demands unbiased decisions, such as automated bank loan application decision-making.
    - Other AI applications do not. We might trust another human's judgment even though the mind is not explainable nor human beings necessarily reliable. For certain applications, then, trust and success do not depend on explainability.
    - In many cases is perhaps more important that a system be reliable – that it works as expected. If it does, an explanation of "how" is less relevant.
- Explainability is a goal but may not always be a reality.
    - In some applications it may be the case that we cannot 'get to explainable.'
    - Explainability should just be one tool in a toolbox alongside others: trustworthiness, effectiveness, and compliance mechanisms. These may include testing, licensures, verification procedures, and other mechanisms.

Panel 2: AI and National Security Ethics: Bias, Data, and Principles

A second panel addressed ethical issues posed by using AI in the national security and especially intelligence contexts. The panelists discussed how to foster the development of AI that will be used in ways consistent with equality, justice, privacy, civil liberties, and human rights. They discussed how to encourage not only our government but also other nation states to abide by those ideals. In considering the larger question of how, ethically, to partner humans with machines, the panelists addressed the sub-issues of how to assign human responsibility for AI outcomes and how to mitigate bias in AI applications. They proposed recommendations and strategies for attorneys to adopt and deal with those issues. The panelists emphasized that every AI system requires its own ethical and legal analysis, and that attorneys should be involved and partner with technologists in that case-by-case analysis, from design through, and throughout, deployment.

Key Points and Observations

As the AI Commission has noted, the benefits of AI for the Intelligence Community include uncovering patterns and trends in intelligence collection and processing, automating natural language and video data processing, and analyzing open-source information. Against the backdrop of benefits of AI to the intelligence community, the panelists discussed some of the most important ethical issues that government attorneys should consider. The panel focused on:

- AI as Ideology
- Responsibility
- Bias, Fairness, and Justice
- Privacy and Data Collection
- AI Values and International Alliances and Cooperation
- Takeaways for Attorneys

AI as Ideology

The panel discussed the ideological assumptions in some of the general discussion about AI as well as the values individual AI applications might encode.

- Referencing a recent publication,[ix] one panelist suggested that AI can be understood as an ideology, not just a suite of technologies: AI could, in the minds of some, become autonomous from and eventually replace, not complement, humans and much of human judgment and our humanity. In some ways that ideology resonates with a number of historical ideologies and authoritarianism. We must keep a keen focus on what it means to put humans and human judgement front and center as we use AI.
- We must be vigilant about what values we instill in AI: we design AI using a value set; in turn, AI uses, enforces, and proliferates those values. It is essential we select and guard these values with care.

Responsibility

As one panelist explained, often when we use AI, we delegate power or a decision from a person to a machine-driven process. Even when a human is in the loop, if AI systems are providing the human with input and guidance, then those machines are partners in the decision-making process. That partnered decision-making is fundamental to reaping the advantages of AI for government and private sector missions. But we must think carefully not only about what we are delegating but also how we assign responsibility, especially because AI creates layers of indirection and disconnection between human decision makers and practical outcomes.

- *Human-Machine Teaming*. One panelist explained how human-machine decision-making can be improved during real time testing in the field by constantly tweaking and adopting algorithms. For example, developers might tweak a semiautonomous, unmanned vessel's navigation system in response to mistakes it makes in test deployments. However, as we employ more and more autonomous systems, it will become increasingly difficult to dedicate time and resources to refining the decision-making of each of those systems. In other words, with the proliferation of autonomous systems, we may be less likely to engage in the type of meaningful human-machine teaming that ethical deployment would require.
- *Aligning Responsibility with Authority and Information.* In considering how to assign human accountability for AI outcomes, panelists suggested drawing lessons from society's experience with issues of responsibility and process in large organizations. In many AI circumstances, there may be shared responsibility across groups, but as we have seen in other contexts, that might result in no one person or office taking responsibility for failures. It is important that those individuals with the opportunity and understanding to check, adjust, or stop AI systems must also have both the authority and information necessary to do so. Responsibility must be aligned with the authority and information necessary to operate the system safely and ethically. It is also important to assign responsibility for building guard rails and processes that will help humans see when things are going wrong so that we can intervene and manage the consequences.
- *Responsibility for Bias.* Panelists suggested that part of taking responsibility for AI includes involving stakeholders, in all stages of development and deployment of an AI system, and necessarily includes mitigating against and monitoring for bias. Any AI system will use data

inputs. As one panelist explained, the "specific snippets of the world that you get from your data will create bias," and as we delegate responsibility to unmanned systems that inform us about the world, through their biased lenses, we must consider how we share responsibility for those individual and collective visions.

Bias, Fairness, and Justice.

Panelists explained how various forms of algorithmic bias can impact AI applications, from instances where an application is too brittle to adjust when data inputs in real conditions differ from those in training conditions, to AI outcomes that reflect human biases, where an application or machine replicates or exacerbates existing discriminatory practices, such as racialized or gendered practices. Algorithmic bias raises ethical concerns especially when it overlaps with (and recreates or even amplifies) human discriminatory biases. The panel first discussed examples of algorithmic bias and then made recommendations about how to mitigate it.

- *Algorithmic bias generally (how it happens).* Algorithmic bias might occur when an algorithm learning from the data notes correlations that the developer was not aware of at first, due to the complexity of the data. For example, a developer might train a computer vision algorithm on images of cars and trucks in a variety of settings, such as cars on dirt roads, trucks in grass, and box trucks on a dark paved road at night; using the data, however, the algorithm might then misidentify an air conditioning unit on top of a dark black building as a box truck. Bias can also stem from a *lack* of data. For instance, a developer trains a UAV in the desert and the UAV is unprepared (lacking in correlations) to operate in a jungle.
- *Biases with disparate social harms.* One panelist suggested demographic disparities in the accuracy of facial recognition algorithms as a well-known example of algorithmic bias with social harms. A 2019 National Institute of Standards and Technology (NIST) study[x] concluded that the error rates of facial recognition algorithms differ across racial and ethnic groups, as well as by age and by gender, and that they differ in ways that disfavor the same groups that often face discrimination elsewhere in society. Consequences to an individual disfavored by the algorithm might include a person receiving undue attention from law enforcement, or that a government record is created about what she has or has not done because she has been mistaken for someone else, or that she is denied some benefit because of such a mistake. Some of the ways social disparities might be encoded in AI are:
    - Unrepresentative data*:* a group might be underrepresented in the data used to train the algorithm.
    - Biased historical data: some systems are trained to try to emulate past human decisions deemed to be correct; if those past human decisions encode human bias, the algorithm learns to replicate that bias. Predictive algorithms used in the criminal justice system, for example, have been criticized for reinforcing historical injustices and disparities.

  - A human developer might encode implicit or explicit bias, intentionally or not, via linguistic and other assumptions.
- *Mitigating Bias*. Panelists recommended that to mitigate bias, government officers and attorneys should:
  - Think intentionally about bias from the beginning of a project instead of addressing it at the end.
  - Consult with the affected communities about bias from the very beginning. Any plan that will affect citizens or particular people in particular categories ought to include those communities.
  - Ensure that all other stakeholders are represented at the requirements and development phase to properly assess the detail needed to target potential bias. In addition to the affected communities, these stakeholders include, among others, the designers, the developers, the regulatory team, the maintenance team, the lawyers.
  - Ensure that the requirements language is precise and accurate so that the right people might be consulted about potential biases.
  - Build in consultation and iterations with stakeholders. One panelist suggested the DOD cyber risk management framework as a helpful analogy.
  - Measure the effects of a system. Develop and use metrics to assess the impact of bias on communities, such as how often people in different groups get a bad result from your system.
  - Continue to measure the effects of a system once it is in operation. Understand that the conditions will change, and the metrics should be adjusted to evaluate the change. AI is “brittle”: when a system trained in one setting is used in another setting, it can go wrong. One of the ways a bias disaster might occur is if conditions change in ways developers or users do not plan for.
  - Work with the technology team to identify whether anti-bias technologies or methods are available based on the specific AI system and setting.
- *Bias in Classified AI*. One panelist noted that while there might be a tendency to consider bias a domestic issue outside the purview of national security lawyers, national security lawyers working in a classified environment have a heightened responsibility to be exceptionally conscious of bias in military as well as civil and human rights contexts.

Privacy and Data Collection

The panel touched briefly on select privacy issues:

- *Data collection and hoarding.* AI technology tends to drive institutions and organizations toward collecting and hoarding data. There exists a mentality that an organization can never have too much data and that data can never be too personal, no matter the comparative value of the organization’s mission. Not only is the collection potentially problematic but so too the “attractive nuisance” created by its retention.
- *Privacy-enhancing technical measures.* These measures within AI are largely context-specific, but these promising new technologies draw inferences about populations in ways

that preserve individual privacy. AI users should consult with their technical teams to understand whether there exist any privacy-enhancing technologies that might work with their AI systems to protect individual privacy.

AI Values and International Alliances and Cooperation

The use of AI technology by authoritarian states and in the context of great power competition presents challenges to the United States and the international community.

- One panelist spoke to policy momentum around cooperation among the United States and its allies to infuse AI norms with liberal democratic values such as liberty, justice, and equality. Allies might pool resources and leverage safe and reliable AI in support of inclusive growth and human rights.
- Areas for alliance include sharing resources and potentially data sets, working together on privacy-preserving technology, promoting interoperability, and shaping norms and technical standards; they also include defensive measures such as preventing the transfer of sensitive technological information and coordinating on investment screening procedures, and exploiting hardware choke points.[xi]
- Despite the dynamics of great power competition, there are areas of commonality between competing nations; the hard work in cooperating will be moving beyond facial agreement to dealing with details at the brass-tacks level, where ethics and values will be at stake. One area of particular concern might be to establish an agreement to maintain human control over nuclear control systems, which the United States has indicated is a core commitment. Negotiations about AI ethics and norms will need to be on a case-by-case, scenario-by-scenario basis to be meaningful. Developing a broad consensus among allies first will help in shaping those negotiations.

Takeaways for Attorneys

The panel concluded with a discussion of the takeaways for national security attorneys:

- Learning the “language of AI” is essential to understanding it. Lawyers should learn what it means, or might mean, when someone says ‘AI,’ ‘automation,’ or ‘autonomous agent’; and understand, too, that a particular system might be comprised of multiple AI (and non-AI) parts and algorithms: a computer vision algorithm, a navigation component, a human operator.
- Lawyers should work in partnership with technical teams to address ethical and policy issues, rather than simply setting out ethical and policy goals and turning projects over to technical teams to build. Technical teams have tools that can help in most legal and policy goals, from compliance to monitoring.
- Engaging with issues surrounding AI technology throughout its lifecycle is important. AI often fails when conditions change, and conditions will change in the national security world. As conditions change, AI systems must be constantly reevaluated for their use in missions and areas of interest, and to ensure they are not going off the rails.
- Untangling bias from and instilling fairness in AI is difficult, perhaps impossible, but it is a battle we can continue to fight and an area in which we can continue to do better.

- Lawyers should have robust conversations with engineers and developers to disentangle the difficulty of explaining what happens inside the black box from the intent of and the risks associated with an AI application. Lawyers are well-trained as a generic matter in being dogged in asking what the risk is to the government or agency mission and our values; it can be harder to press for details and answers on complicated technical matters, but it is critical to do so.

Panel 3: AI and National Security Decision-Making

Much of the national security debate about AI law and ethics centers on certain applications, such as lethal autonomous weapons, and on public and private bulk data collection, control, and management. Bias is also a central theme in AI discussion. This includes the many ways mentioned in the previous panel that witting and unwitting bias can impact the design, use, and predictive accuracy of algorithms.

The role of AI in informing, augmenting, or executing national security decisions warrants equal ethical attention. There are many potential and realized AI applications that should augment human decision-making capacity. A technological capability that can help find and correlate data into discernible intelligence information and do so at machine speed should, for example, in theory, improve decision-making outcomings. Likewise, if an AI driven computer can be taught, or teach itself, to play chess or GO by predicting its opponent's moves, AI can be designed to predict and model foreign policy and national security outcomes. But there are also challenges and risks. Predictive algorithms, for example, can facilitate decision-making by augmenting the human capacity to assemble and find meaning in data, but they can also embed existing societal and other biases, as debates about algorithms that predict parole and bail risks suggest. We also know from cyberspace some of the risks of operating at machine speed, such as the need to pre-delegate responses. There are also heightened and, in some cases, unique challenges and risks, such as those associated with attribution, security, and collateral effects when identifying and verifying software code.

We should want decision-makers to use all available tools to inform and make decisions, but we should also want them to do so wisely, aware of the risks, benefits, and limitations of the tools they are using. One risk is that decision-makers will not fully appreciate or understand the tools they are using, including what is occurring in the black box of deep learning neural networks. As the Symposium explored, there are mitigation measures that can address some of the decision-making risks of AI. These include, testing, training, and making sure the right people are in the decision-making room when AI is used. The third panel addressed decision-making exploring these issues in detail highlighting five areas:

- The State of Government-Developed AI
- Law and Policy Gaps, including in Accountability
- Human Machine Teaming
- Building Ethics and Security into Development
- The Role of Lawyers

The State of Government-Developed AI

The panel first addressed some challenges and successes DOD has had in developing AI:

- Significant parts of DOD are operating on a very different information technology infrastructure than commercial entities. For example, DOD is still migrating to a cloud-based infrastructure, and much of DOD is not operating on the cloud. Similarly, much of DOD does not have a continuous integration pipeline to consistently update software. While the private sector may update an operating code within a couple weeks, parts of DOD may update software once every seven years, if at all.
- That said, there are “islands of incredible performance” within DOD, “delivering some really exciting applications” in prototype or in operation. One such application is predictive maintenance software for military helicopters. This software synthesizes and integrates data for each aircraft part and uses machine learning to predict which helicopter is likely to experience maintenance failures. Using this software, the government can anticipate mechanical issues and fix them prior to flight, reducing the costs of aircraft maintenance and risks to personnel.
- Further, the DOD Joint Artificial Intelligence Center (JAIC) is executing dozens of AI projects across fields as disparate as warfighter health, humanitarian assistance and disaster relief, and joint war fighting in traditional combat. As these projects underscore, there is a gap between the public debate about AI in the national security domain and how AI is actually being used. The debate tends to focus on “unicorn technologies,” such as lethal autonomous weapons systems with the potential for human casualty. While those technologies do present significant risks, much of what the government is actually developing is intended for very different contexts.

Law and Policy Gaps, including in Accountability

A majority of the panel took the view that law and policy fail to keep pace with technological development.

- One place where current law and policy is underdeveloped is with respect to accountability for AI.
    - One panelist argued there is undue focus on criminal liability as an appropriate response to artificial intelligence accidents. Instead of placing emphasis on criminal liability, we should be thinking of accountability and responsibility in broader terms. We should begin implementing administrative accountability mechanisms, which are predominantly used in the government. Those mechanisms do not focus on criminal penalties, nor require a showing of intent, but rather look at organizational accountability and learning how to correct failures in organizational decision

making, both retrospectively and moving forward. Administrative accountability mechanisms might include: commander's investigations, Army 15-6 investigations, commissions of inquiry, advisory committee task forces, and Inspector General's investigations.

  - The benefits and challenges of administrative principles should be discussed further. We should focus on how to make these processes more effective, independent, and impartial, with the right mix of backward- and forward-looking remedies when problems arise.

- One panelist, however suggested that existing law and regulations should not be under-estimated in their current capacity to effectively regulate today:
  - While AI and technological development in general drive great change, that is not new to DOD. People outside government may underestimate the quality of existing regulations and processes and their relevance to any new technology. For example, the LOAC principles – necessity, proportionality, humanity, distinction, and honor – have been in place for many years. The DOD *Law of War Manual* is very long; we have thought long and hard about those principles.
  - The panel tended to agree that DOD's ethical standards have risen as technology has become more capable. Precision-guided munitions provide an example. Although the *Terminator* movies of the 1980s were written in response to, or in fear of, heat-seeking, precision missiles, we cannot imagine a situation today where a military commander would use anything other than a precision-guided munition in an urban area. Militaries all over the world in the 1940s viewed carpet bombing as an acceptable use of force; today, the U.S. military in particular would never allow the use of anything other than a highly precise munition in that kind of environment.
  - In addition to LOAC, regulations and other mechanisms for compliance are both important and well-developed. A three-million-person organization such as DOD faces the risk that a thousand-to-one or one-in-a-million type problems will occur. Regulations and compliance mechanisms help DOD to know and show that it is upholding the law. DOD has for many decades tried to ensure that such a large organization can still uphold ethical obligations.
    - DOD already has an excellent body of regulations and processes that can be applied to artificial intelligence instead of developing an entirely new regulation process. These existing mechanisms for compliance are demanding. For example, Military Standard 882E[xii] governs system safety of all major defense acquisition programs. If the operation of a system might have life or death consequences, whether it is an aircraft that might crash or a system that uses force, 882E requires the government to present technical evidence suggesting that the chance of a software failure is less than one in ten million. As a result, the panelist suggested, there is not a chance a "cigar-chomping general" is going to "send in the killbots" or field "some crazy half-baked, undertested system."

When asked about the mismatch between the pace of technological change and the pace of law and policy development, one panelist responded that:

- Law lags far behind technology. The United States has moved toward developing and acting based on policies that may go above and beyond the requirements of the law.
    - A policy framework can be a valuable approach: it is flexible, and can be adapted quickly, unlike, for example, international law, which is slow to develop.
    - But it can be confusing where it is not clear where the line between what is law and what is policy begins and ends. That is something specialists and lawyers should pay more attention to, both from perspective of states and other actors.
- With respect to the "very, very high standards the U.S. military follows," there is a pressure created by the development of the precision capability that pushes countries like the United States above and beyond what the law requires. The U.S. might, as a matter of policy, go over and above what the proportionately rule or what feasible precautions might require. Some groups will then argue the U.S. is legally required to use the highest precision system available.
    - The panelists agreed that this creates a challenging dynamic, and, problematically, a disincentive for the government to follow the highest possible ethical standards.
- The blurred line between law and policy should be demarcated. From the government perspective, and from a civil society perspective, if the line is not drawn, then there will be pressure from outside groups that will argue the law is constraining how the government can operate while in reality it could be policy and ethics that is constraining government action, not law. Thus, it is important for lawyers to identify what legally the government must do or can do. Without such a line, government actors may be disincentivized from applying higher ethical standards as a matter of discretion that might otherwise be viewed, properly or not, as legally constraining future decisions. By bringing clarity to the line between law and policy and ethics, we might see increased development of artificial intelligence ethical standards.
    - To begin identifying this line, the government should make clear when it is acting pursuant to higher ethical principles and distinguish those principles from what the law actually requires. Civil society groups should recognize that there are risks to arguing that practice means the law has changed because it can create disincentives.

Human Machine Teaming

The panelists next discussed how we might see law and policy change in response to human machine teaming.

- Human machine teaming is not new and can provide great benefits. One example from the 1980s: the F16 fighter jet system automatically takes control over the aircraft when it senses the pilot has lost consciousness.
- A significant challenge for human-machine teamed systems is pinpointing individual responsibility when failure occurs. One panelist argued that when failures do occur,

determining who is criminally responsible is the wrong approach. Instead, we should look at administrative principles, as described above.

- One panelist noted that success in human machine teaming depends on identifying AI "failure modes." The challenge will be identifying what these failures look like, testing for those failures, developing methods to continuously monitor for these failures, and thereby producing reliable systems that work in a diverse set of environments.
- Panelists noted the fine line between seeking the benefits of teaming and accepting certain risks of greater AI autonomy. With greater autonomy comes greater risk of machine failure and potential catastrophic machine trophic failures. The more autonomous control a system has, the more harm that could be done and the greater the risk of unintended escalation.
    - Machine learning allows us to envision a much longer period of automation over a more diverse set of environments. A heat-seeking missile makes its own decisions after launch, absent an abort switch; newer systems will have longer and more varied periods of automation, increasing the risk for failure.
    - Existing DOD regulations are very strict, and DOD is still researching what are the best methods to allow for confidence in our testing. Verification and validation is a huge area of ongoing research.

Building Ethics and Security into Development

The panel discussed DOD's efforts to consider ethics and security at every stage of AI development and use.

- Earlier this year, DOD issued its ethical principles for AI, demonstrating a commitment to ethics.
- The panel emphasized the importance of incorporating ethics into every stage of AI software development, as suggested by the informally named DOD practice: "DefSecEthOps."
    - "DevOps" is a management practice adopted from Silicon Valley to encourage agility in technology management. It merges software development and IT operations (upon which the software, here AI, runs). It is an approach designed to mitigate against the cliché situation where an organization is constantly fixing what broke during the last round of software updates. That is particularly important when a large organization wants to change thousands or millions of machines by fielding software updates and those updates might have life or death consequences: DOD must be extremely sure its software operates exactly as intended.
    - DOD has built "Security" into its DevOps approach: "DevSecOps." Given the increasing demands of cybersecurity, and the constant attacks on the government in particular, developing and implementing new software must be done in a way that emphasizes security at every step.
    - Likewise, ethics must be emphasized at every stage of software and, in our case, AI development and lifecycle: hence the term, "DevSecEthOps." Ethical failures can occur at any point of an AI software program: for example, if developers obtain or use data in a way that violates someone's privacy or if commanders deploy an application in a situation where it is inappropriate and compromises the safety of

the operator or someone effected by the system. We need to be conscious of the ethical challenges we are likely to encounter throughout the course of our work and to develop best practices to stay on the right side of ethical lines.

The Role of Lawyers

The panel emphasized the important and evolving role of lawyers in AI decision making.

- One panelist highlighted lawyer productivity as a key focus of the AI Age. Workflows tend to bottle neck with lawyers (or so it was alleged!). DOD is currently working on developing natural language processing AI to help lawyers research and identify relevant law and policy from the corpus of DOD regulations. Such software would go beyond keywords, instead using a semantic understanding of words to allow lawyers to research with plain language queries and receive nuanced answers. For instance, a lawyer could potentially ask, in natural language, 'if we want to do x, can we do that? whom do I need approval from?' This type of application would be first made available for the legal and policy community of DOD, because we know that to be a nation ruled by laws, people should have a reasonably rapid ability to understand what the law is, how it is being implemented, and by what policies. In essence, DOD/JAIC is developing human machine teaming for lawyers.
- Lawyers will play a part in development as the representatives of the ethics ingredient in DevSecEthOps. If at every stage in development, developers design alongside lawyers and ethicists, decisions will better account for ethics and the law, minimizing blind spots.

Keynote Session: Lawyers and Law

If AI will transform national security practice, the question for lawyers now becomes how will AI transform national security legal practice? How should national security lawyers, and all lawyers, contribute to the AI mission? The question starts with an understanding of the three purposes of national security law: (1) The authority to act, along with the boundaries of that action; (2) the provision of essential and effective process; and (3) the sustainment of our core legal and constitutional values, which in many cases also reflect core national security values. Symposium participants made the following points.

Law and ethics will, or could and should, distinguish democratic and American AI from authoritarian applications of AI.

- Adherence to American concepts of law and ethics will, for example, help determine whether and to what extent U.S. companies and talent help authoritarian regimes govern and control their populations.
- The ethical use of AI is also more likely to attract and retain AI talent to national security missions and result in stronger public-private-academic partnerships.
- Law and ethics will bind like-minded alliances in the AI field, and they will help to build and sustain public trust and support for appropriate AI applications. The converse is also likely. If, for example, the public does not trust the government's use of AI in one context, for example, because of certain facial recognition applications, it may not trust the government with using AI in other contexts, for example, to facilitate contact tracing amidst a pandemic.

Lawyers should know that they can make critical contributions to technology policy.

Part of the reason CSET exists is to bring together thought leaders to learn more about law and policy. As one panelist noted, "I had spent a decade on technology development and I was pretty ignorant of technology policy; and if I had it to do all over again, I'd spend much more of my time getting smarter on law and policy before I joined government." Lawyers who have an interest in technology are, collectively, a precious resource, as are technologists who have an interest in law and policy. Every lawyer should make friends with a technologist, and every technologist should make friends with a lawyer.

Therefore, lawyers need to understand how AI is going to change the nature of national security legal practice and potentially do so in profound ways. For example, AI will impact many, if not most of the fields of practice: logistics, intelligence, decision-making, communications integration, cybersecurity, information processing, how we classify and declassify information, hiring practices and so on. AI will also change the speed of decision, or can, when we allow it to or let it do so, as has happened in cyber space. Lawyers, as counselors, will need to help policymakers and technologists, in government and out, figure out where and when to rely on algorithms to inform decision, make decisions, or augment human capacity. In the national security space, as in cyberspace, we will need to sort out when and how to pre-delegate decisions or decision-making authority and actions. We already have experience doing this in cyberspace and in certain military command and control contexts. For example, we need to consider what lessons, if any, we may have learned and developed devising nuclear doctrine and principles that might apply or be adopted to AI applications. One lesson is that we do not always get the doctrine right at the outset.

Lawyers will also have to change how they practice.

With AI, this means moving upstream into the research and development stages of AI rather than waiting downstream at the point of use. As lawyers already know from policy practice, if you wait "to lawyer" at the use or decision points, it may be too late to meaningfully influence outcomes as policymakers are locked into choices already made. With AI it may be too late to fully understand the potential risks and impact of issues like data and design bias. Lawyers must be ready throughout the AI lifecycle – upstream and downstream - to advise on AI use, maintenance, and adjustments, to the end.

AI will also change the nature of the questions lawyers should ask. (One of our panelists has offered lists of questions lawyers should ask about data, bias, and algorithms in the book *The Centaur's Dilemma*. The questions are intended as a place to start for lawyers to engage technologists on the design and use of AI.) The bottom line is that it is time for national security legal practitioners to move from bromides about "humans in the loop" and "AI principles" to the specific application of those principles to unique AI applications. In doing so, lawyers will need to recall that they are not just bringing the substance of the law as it exists to the process, but all the considerations contemplated by Model Rule 2.1, like ethics, making it clear when they are applying hard law, good judgment, or simple common sense.

Evolving Law: Use boundaries; data; redlines; and predictive algorithms.

Lawyers will debate the use limits of AI, such as the use of predictive algorithms, as they already are regarding parole and bail algorithms. And lawyers will debate use redlines around such applications as autonomous weapons and deep fakes. Again, they already are. However, the most significant impact from lawyers and the law may come in less visible areas of practice. The most valuable input will probably occur in the back-office process of automation (finance, contracts, logistics) because that is where we spend so much time and money. And while we work out some unsolved problems in the security and safety of AI systems, that does not need to slow down the application to the back-office applications, where there is not so large of an attack surface.

Lawyers will also likely play an oversized role in helping to design and implement the design or architecture of how the government goes about making AI decisions, whether those decisions occur in the R & D (JAIC) stage or use stage of AI development. We are getting better. The creation of the Joint AI Center was probably the most significant development to date. We now need to elevate its Director higher in the org chart. We also need to empower other departments by including similar or model capacities: there is a proposed Assistant Secretary of State for Emerging Tech, to lead on tech diplomacy. The government also needs an intelligence cell focused on unclassified global science and technology developments, greater White House coordination on technology across the National Security Council, Office of Science and Technology Policy, and National Economic Council, and increased investments in the National Institute of Standards and Technology to lead on technology standards. Many of these are recommendations from the NSCAI.

Finally, if policymakers are going to wield the benefits of AI and do so wisely, it matters who is in the room. In creating good process, lawyers will have an opportunity to shape outcomes by making sure the right people are in the room when it comes to designing and using AI. At the national level, this may mean having more junior officials attend senior meetings to explain AI generated intelligence or predictive modeling along with its predictive strengths and limitations.

The panelists concluded by noting that the time is now for improved teaming between lawyers, policymakers, and technologists and to make purposeful legal and ethical choices in how AI is developed and used.

Conclusion: The Centaur’s Choice

In U.S. national security practice, specialists refer to the relationship between AI and humans as human-machine teaming. At the Department of Defense this process is sometimes referred to as a centaur model of AI decision-making and use, part human and part machine. The challenge it turns out is often in determining when and how to effectuate this teaming so that decision-makers can reap the benefit of the AI application in use, and do so at machine speed, without losing control of or an understanding of the outcome. (Hence the terminology “human on, in, or out of the loop.”) Depending on whom one is talking with this process of teaming is either well in hand – safe, secure, and steady – or it is nascent and fragile.

Our symposium did not seek to reach a conclusion on this or other points of discussion. Our goal was to remind the audience that we have a choice, let us call it the centaur’s choice, in how we shape the national security uses of AI and how we address, and hopefully mitigate, the risks of using AI applications in the security field. We are not passive actors. Machines do what they are designed to do or programmed to do. Not all of us may know yet where we are headed with AI, but we do know we are certain to get there. Therefore, national security lawyers need to become AI generalists and team with policymakers and technologists in the development and use of AI. Conversely, policymakers and technologists need to understand the law, so that they can spot issues and create the time and space to embed ethical and legal principles in AI applications, not just to comply with the law, but to ensure AI is used more effectively and accurately.

While our panelists expressed different and, in some cases, varying views on the trajectory, ethics, and law associated with AI, there was agreement on this: U.S. national security will be better served with the meaningful, thoughtful, and purposeful application of law and ethics to the use of AI. We hope the Symposium and this report will help in some small way contribute to this result.

Symposium Agenda and Participants

Optional Introduction to AI

Matthew G. Mittelsteadt
Artificial Intelligence Policy Fellow, Syracuse University Institute for Security Policy and Law

Introduction

James E. Baker
Director, Syracuse University Institute for Security Policy and Law

Panel 1: AI & the Law of Armed Conflict

Margarita Konaev
Research Fellow, Georgetown University Center for Security and Emerging Technology

Jason R. Stack
Director, Ocean, Atmosphere, and Space Research Division, US Office of Naval Research

Iben Yde
Assistant Professor, Royal Danish Defense College; Former Legal Advisor to the Admiral Danish Fleet and Joint Defence Command, Denmark

John R. Cherry, Moderator
Deputy Chair and Military Professor, Stockton Center for International Law, US Naval War College

Panel 2: AI & National Security Ethics: Bias, Data, & Principles

Tarun Chhabra
Senior Fellow, Georgetown University Center for Security and Emerging Technology; Nonresident Fellow, Brookings Institution

Andre Douglas
Project Manager and Section Supervisor, The Johns Hopkins University Applied Physics Lab

Edward W. Felten
Board Member, Privacy and Civil Liberties Oversight Board; Director, Princeton University Center for Information Technology Policy

Laurie N. Hobart, Moderator

Assistant Teaching Professor, Syracuse University College of Law

Panel 3: AI & National Security Decision-Making

Gregory C. Allen
Chief of Strategy and Communications, US Department of Defense Joint Artificial Intelligence Center

Laura A. Dickinson
Oswald Symister Colclough Research Professor of Law, The George Washington University Law School

Reginald Brothers, Moderator
CEO, NuWave Solutions; former Under Secretary for Science and Technology, US Department of Homeland Security

Keynote Session: Where Are We Headed?

Jason Matheny
Founding Director, Georgetown University Center for Security and Emerging Technology; Commissioner, National Security Commission on Artificial Intelligence

James E. Baker
Director, Syracuse University Institute for Security Policy and Law

Acknowledgments

The Syracuse Institute for Security Policy and Law thanks Jason Matheny and the Georgetown Center for Security and Emerging Technology for its support in hosting this symposium. SPL also would like to sincerely thank Kristen Duda, John Cherry, Martin Walls, Kyle Davis, and research assistants Hannah Gabbard, Thomas Finnigan III, and Aly Kozma for their assistance in organizing the symposium, designing the program, and facilitating the registration and platform, all done with care and humor. This report was drafted by Jamie Baker, Laurie Hobart, Matt Mittelsteadt, John Cherry, with contributions from Hannah Gabbard, Thomas Finnigan III, and Aly Kozma.

Special thanks also go to our panelists and moderators who did a wonderful job framing questions and sustaining an engaging and iterative dialogue over three hours of remote engagement, as opposed to offering sequential speeches. Finally, we would like to thank our remote audience of over 180 participants for sharing the afternoon with us and (we hope) agreeing to jump into the subject of AI so that we can better maximize the benefits and mitigate the risks of the important and emerging national security tools enabled by AI and machine learning with the input of law, ethics, and good process. Thank you to all.

Endnotes

i Interim Report, National Security Commission on Artificial Intelligence, 9, November 2019, https://www.nscai.gov/wp-content/uploads/2021/01/NSCAI-Interim-Report-for-Congress_201911.pdf.

ii Greg Allen and Taniel Chan, “Artificial Intelligence and National Security,” 1, Belfer Center for Science and International Affairs, Harvard Kennedy School, July 2017, https://www.belfercenter.org/publication/artificial-intelligence-and-national-security.

iii Cate Cadell and Adam Jourdan, “China Aims to Become World Leader in AI, Challenges U.S. Dominance,” *Reuters*, July 21, 2017, https://www.reuters.com/article/us-china-ai-idUSKBN1A5103.

iv Interim Report, National Security Commission on Artificial Intelligence, *supra* note 1.

v Paul Scharre, “Autonomy, ‘Killer Robots,’ and Human Control in the Use of Force – Part I,” *Just Security*, July 9, 2014, https://www.justsecurity.org/12708/autonomy-killer-robots-human-control-force-part/.

vi Department of Defense Directive 3000.09, “Autonomy in Weapon Systems,” updated May 8, 2017, https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf?ver=2019-02-25-104306-377.

vii *Department of Defense Law of War Manual*, Office of the General Counsel, updated December 2016, https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20-%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190.

viii “DOD Adopts Ethical Principles for Artificial Intelligence,” Department of Defense press release, February 24, 2020, https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/.

ix Jason Lanier and Glen Weyl, "AI Is an Ideology, Not a Technology," *Wired*, March 15, 2020, https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/.

x Patrick J. Grother, Mei L. Ngan, and Kayee K. Hanaoka, “Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects,” National Institute of Standards and Technology Interagency/Internal Report (NISTIR) 8280, December 19, 2019, https://www.nist.gov/publications/face-recognition-vendor-test-part-3-demographic-effects.

xi Andrew Imbrie, Ryan Fedasiuk, Catherine Aiken, Tarun Chhabra, and Husanjot Chahal, “Agile Alliances: How the United States and Its Allies Can Deliver a Democratic Way of AI,” Center for Security and Emerging Technologies, 2020, https://cset.georgetown.edu/research/agile-alliances/.

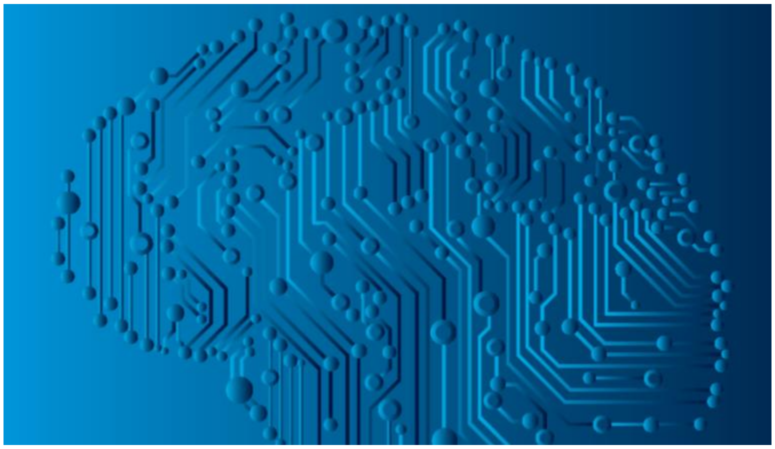
xii Military Standard (MIL-STD) 882E, Department of Defense Standard Practice of System Safety, revised May 2012.

Further Recommended Reading

Reports, National Security Commission on Artificial Intelligence, 2019-2021

The IC Principles of Artificial Intelligence Ethics and the IC Artificial Intelligence Ethics Framework, 2020, https://www.dni.gov/index.php/features/2763-principles-of-artificial-intelligence-ethics-for-the-intelligence-community.

The IC Principles of Artificial Intelligence Ethics and the IC Artificial Intelligence Ethics Framework, 2020, https://www.dni.gov/index.php/features/2763-principles-of-artificial-intelligence-ethics-for-the-intelligence-community.



Syracuse University Institute for Security Policy and Law

300 Dineen Hall
950 Irving Ave.
Syracuse, NY 13244

securitypolicylaw@law.syr.edu

Center for Security and Emerging Technology

Georgetown University Walsh School of Foreign Service
37th St. NW & O St. NW
Washington, DC 20057

cset.georgetown.edu