iStrategy: How AI will Augment Future Strategy Decision-Making

by

Lieutenant Colonel Michael J. Winter
United States Air Force

Under the Direction of:
Dr. Andrew A. Hill

United States Army War College
Class of 2018

DISTRIBUTION STATEMENT: A
Approved for Public Release
Distribution is Unlimited

The views expressed herein are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Army, Department of Defense, or the U.S. Government. The U.S. Army War College is accredited by the Commission on Higher Education of the Middle States Association of Colleges and Schools, an institutional accrediting agency recognized by the U.S. Secretary of Education and the Council for Higher Education Accreditation.
**ABSTRACT**
Modern trends in artificial intelligence (AI) indicate a rapidly approaching intellectual revolution, which implies AI will one day extend human cognition and augment human strategy decision-making. Strategic leaders must now contemplate the future of strategy decision-making through three paradigms—human-only, human-machine, and machine-only. The U.S. can obtain a competitive advantage in strategic-level cognitive power—and thus judgement superiority—by leveraging the human-machine paradigm to generate “iStrategies” to complement existing Department of Defense offset proposals. A Department of Defense strategy on AI is required to not only address challenges to the AI-enabled paradigms, but to also meet current national-level guidance and intent on AI.
Abstract

Modern trends in artificial intelligence (AI) indicate a rapidly approaching intellectual revolution, which implies AI will one day extend human cognition and augment human strategy decision-making. Strategic leaders must now contemplate the future of strategy decision-making through three paradigms—human-only, human-machine, and machine-only. The U.S. can obtain a competitive advantage in strategic-level cognitive power—and thus judgement superiority—by leveraging the human-machine paradigm to generate “iStrategies” to complement existing Department of Defense offset proposals. A Department of Defense strategy on AI is required to not only address challenges to the AI-enabled paradigms, but to also meet current national-level guidance and intent on AI.
iStrategy: How AI will Augment Future Strategy Decision-Making

“The most powerful actors of the future will be the states, groups, and individuals who can leverage...relationships and information in a more rapid, integrated, and adaptive mode than in generations past.”

—National Intelligence Council

Modern trends in artificial intelligence (AI) indicate a rapidly approaching intellectual revolution, one where digital technology will elevate human cognition to levels analogous to the impact the steam engine had for human muscular strength during the industrial revolution. While historically humans have dominated strategy decision-making, the current technological trajectory of AI warrants a reevaluation of that monopoly. Furthermore, a reevaluation is especially relevant given our cognitive limitations to operate in a strategic environment that the Department of Defense assesses will be “more complex and volatile as...conflicts become increasingly transregional, multi-domain, and multi-functional.”

Thinking strategically about AI today is similar to strategic thinking on cyberspace in the mid-1990s prior to the establishment of cyberspace as warfighting domain in 2010. However, current defense strategic thinking and debate on AI—from the Defense Innovation Initiative to the conception of new Department of Defense offset strategies—focuses strictly on the operational and tactical-level use (or non-use) of automated and autonomous systems. Therefore, the Department now needs to be strategic in thinking about future human relationships with computers and how to better use data and information. A key strategic opportunity in the future will be leveraging technological advances in AI that will learn and reason under uncertainty to one day extend human cognition and augment human strategic-level decision-making.
This paper first provides a general framework for how a strategic leader can contemplate future strategy decision-making through three paradigms: 1) a human-only paradigm that provides purpose; 2) a human-machine teaming paradigm that generates concepts; and 3) a machine-only paradigm focused on tactical-level employment. Second, this paper will argue that the human-machine teaming paradigm would best leverage AI to significantly augment (not replace\(^5\)) human cognition and subsequent strategy decision-making to enable strategic-level cognitive power—and transition from present-day information superiority to next-generation judgment superiority. Therefore, this paradigm will one day lead to a U.S. competitive advantage in national-level decision-making exemplified through the human-machine generation of “iStrategies.” Third, this paper will highlight the present-day challenges to a human-machine teaming concept. Finally, this paper will offer recommendations for a Department of Defense strategy on AI to realize the AI-enabled paradigms for strategy decision-making.

**On Artificial Intelligence**

In the iconic science fiction film *2001: A Space Odyssey* (1968), sentient spaceship crewmember HAL (or the Heuristically programmed ALgorithmic computer) goes rogue against the *Discovery One* crew and directly threatens astronaut Dave Bowman by saying, “I know that you and Frank were planning to disconnect me, and I’m afraid that’s something I cannot allow to happen.”\(^6\) While this science-fiction portrayal of AI is far (and very questionable) from ever being realistic, it does serve as a very useful parable to 1) introduce a more prudent definition of AI; 2) frame the more likely technological trajectory of AI in the near future; and 3) provide a framework of how AI will soon replicate human decision-making.
A common definition of AI typically contains two distinct but reinforcing aspects, as it refers to “the scientific understanding of the mechanisms underlying thought and intelligent behavior, and the embodiment of these principles in machines that can deliver value to people and society.” The first aspect of the definition of AI deals with characterizing intelligent performance and behavior, which must also include a social element of intelligence that is oftentimes overlooked by the AI community. The second aspect deals with an entity or system that performs tasks that normally require human intelligence, for example, decision-making. AI is now framed in terms of what humans want to achieve with it, as AI is now generally considered feasible if a human can already perform the function, task, or behavior.

The AI research field, industry, and markets have blossomed in the last decade due to a number of technological and talent advancements that have triggered a recent “second wave” of AI. Six decades of exponential growth in computing power led to a compelling surge in computing performance. Additionally, large volumes and diverse variety of “big” data through cloud-based networked systems are now available for integration with AI systems. Machine learning—or programming a system to recognize patterns and learn inductively from examples and experience—has also flourished with the advent of deep learning neural networks, as machines now loosely mimic human neural functions to learn and reason instead of relying on exhausting amounts of lines of logic-based instructions. Finally, AI computer science and engineering expertise have thrived to meet the surge in commercial investment and talent demand.

Current technological trends in AI now indicate an approaching “third wave”—or advancement beyond the second wave of narrow AI (or task-based) to more general
AI that would exhibit intelligent behavior at least as advanced as a person and perform a full range of complex cognitive tasks.\textsuperscript{12} Advancements in nanotechnology indicate not only a continuance of Moore’s Law (doubling of transistors on a silicon wafer every 24 months), but the possibility that computing power could surge at a future exponential rate.\textsuperscript{13} Next, the volume and variety of available data—stimulated by the likely possibility of \textit{trillions} of networked computers\textsuperscript{14}—is now projected to double every two years, “meaning that as much data will be created in the next 24 months as over the entire prior history of humanity.”\textsuperscript{15} Additionally, AI is now in the early stages of evolving from human-supervised machine learning to active, unsupervised machine learning with the potential to scale a machine’s ability to generalize across a spectrum of task domains and from past experience.\textsuperscript{16}

Therefore, the third wave of AI implies that machines in the near future will fully \textit{integrate} the abilities to see, hear, and reason\textsuperscript{17} in environments with high degrees of ambiguity and complexity\textsuperscript{18} (Figure 1). Deductive, abductive, and inductive forms of reasoning\textsuperscript{19} will soon be integrated across AI algorithms, machine learning, and deep learning neural networks, which increases the potential for better than human performance in many intellectual and \textit{decision-making} behaviors. First, AI will evolve an unsurpassed ability to identify patterns and make characterizations,\textsuperscript{20} as machines will rapidly distinguish across broad data sets, discover connections, and determine similarities and dissimilarities. Second, AI will possess the general ability to conduct data and root cause analysis, as AI will explore the underlying
mechanisms across the data to understand the root causes that drive complex issues and situations. Third, AI will conduct forecasting capabilities, applying predictive analytics to the insights discovered through pattern identifications and causal analysis. And finally, AI will provide prescriptions to complex situations, while evaluating the risk inherent in those recommendations. Thus, as AI continues the trend towards more generalized capabilities, machines will discover more insights, indications, and details from raw data to generate a more comprehensive view of a situation or issues.

Strategy and the Three Paradigms of Future Strategy Decision-Making

Although there are many common definitions of strategy, Basil Liddell-Hart offers the most concise by describing it as the “art of distributing and applying military means to fulfill the ends of policy.” Therefore, strategy is how leadership will use all elements of state power to achieve political objectives. It is also important to highlight the importance of the social feature of strategy. Since warfare is “socially sanctioned violence to achieve a political purpose” according to Joint Publication 1, it thus follows that strategy is socially-constructed given that the decision-making is related to organized violence on behalf of a group.

While strategy formulation and decision-making has historically been human-only, there are now three conceivable paradigms for national strategic decision-making in the coming decades given the trajectory and implications of AI: a human-only paradigm, a human-machine teaming paradigm, and a machine-only paradigm (Figure.)
2). While the three paradigms will be fully integrated in future strategy decision-making across the levels of war, each will bear separate specializations with the human-machine paradigm providing purpose, the human-machine teaming paradigm generating concepts, and the machine-only paradigm focused on employment.

Human-Only Paradigm: Providing Purpose

In *On War*, Clausewitz concluded that war is neither art nor a science, but “rather it is part of a man’s social existence.” Unsurprisingly, humans have dominated strategy decision-making for a number of reasons including superior qualities of intelligence, critical thinking, creativity, and empathy. Compared to other species, humans possess an unmatched “ability to perform highly complex cognitive tasks and behaviors, such as language, abstract thinking, and cultural diversity” due to an advanced human neocortex in the brain. Creativity—which is the highly advanced combination of mental and conceptual abilities—is a key human quality, as the overarching consensus among scientists today is that creativity is a uniquely human trait. Additionally, the human cognition system is significantly bolstered by a social cognition ability in the temporo-parietal junction and prefrontal cortex that helps humans “understand, process, and recall our interactions with others.” Furthermore, humans have been proven to possess superior cognitive capacities and abilities such as functional
knowledge, causal reasoning, foresight, social learning, social intelligence, empathy, and language that directly led to the development and use of complex tools, from primitive cutting tools to modern AI. Thus, humans possess unparalleled cognitive abilities to navigate and operate across complex and abstract environments.

However, humans suffer from a number of physical and psychological limitations that considerably limit human judgment—or “the cognitive aspects of the decision-making process”—and will continue to hinder strategic decision making today and into a more complex and volatile future. Physical limitations that impact human judgment include stress, fatigue, boredom, and hunger. The origin of human psychological limitations involves how the human mind processes information, as humans typically bound rationality during decision-making and satisfice with the most acceptable solution. Humans also suffer from egocentric tendencies that can reject alternative points of view, lead to false assumptions, and warp how information is evaluated. Therefore, human judgment often deviates from rationality since it is regularly biased by numerous cognitive heuristics, which are cognitive shortcuts humans use to simplify decision-making. While typically very useful, these heuristics can frequently lead to biased and inappropriate use that could be potentially harmful in very complex situations, including strategy decision-making in complex adaptive systems characterized by routine volatility, uncertainty, complexity, and ambiguity. Thus, even at our best, human judgement is limited in a multitude of ways.

Furthermore, modern machine systems are now rapidly generating many cognitive strengths over humans including processing power and speed, information volume, and pattern recognition that will continue to mature well beyond human abilities.
In fact, AI can now analyze excessive quantitates and varieties of data at a speed and rate that human biological minds will unlikely ever match.\textsuperscript{38} For example, AI is presently assessed to be 500,000 times faster\textsuperscript{39} at “reading” thanks to modern language processing algorithms, meaning AI is postured to read and synthesize complex findings across the range of scientific and academic journals, articles, books, papers, notes, after-action reports, and lessons learned. Moreover, AI is less susceptible to human evolutionary, biological, and social motivations\textsuperscript{40} to bound reality given the machine ability to process and comprehend an entire spectrum of raw data and information.

Will there still be a role for humans in future strategy decision-making, especially when considering the trajectory of future AI decision-making behaviors? Clausewitz also highlighted that “no one starts a war...without first being clear in his mind what he intends to achieve by that war.”\textsuperscript{41} The human-only paradigm will indeed endure in future strategy decision-making, but will focus primarily on determining and providing purpose—or the ends—to both humans and machines (Figure 3). The most vital strategic-level boundaries of AI are found within the first mile of strategy, including determination of the national interests, priorities, and the political goals, ends, and objectives. Humans alone must always retain national command authority and the decision of going to war (\textit{jus ad bellum}). Additionally, humans must also determine the moral constraints and restrictions on the use of force, as a machine cannot\textsuperscript{42}—and should not—weigh moral considerations and make national-level moral judgments.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Figure 3}
\end{figure}
At the operational level, the first mile boundaries would include translating strategic guidance into military objectives, defining the problem, establishing a commander’s intent and end state, determining termination criteria, and establishing the criteria for the way the war will be conducted (*jus in bello*). Multi-domain command across the operational and tactical levels of war must also remain the domain of humans only.

AI will also be excluded from the last mile of the strategy which includes the evaluation and assessment of whether the resulting strategy realized the political goals and objectives originally set (or dynamically modified) by senior leadership. This starts with the human making an impartial assessment of the acceptability of the ways, the feasibility of the means, and the suitability to accomplish the stated political goals and objectives. Finally, AI must also be excluded from any final evaluation of strategic victory, as any such determination of *status quo ante bellum*—or a “better state of peace”\(^{43}\)—demands a human objective assessment. As such, humans should never yield a final determination of strategic-level victory to a machine.

**Human-Machine Teaming Paradigm: Generating Concepts**

The foundation of the second paradigm of strategy decision-making stems from what is referred to as “Moravec’s paradox,” which states that computers do well what humans do poorly, and vice versa. The human-machine teaming paradigm combines human cognitive strengths from the first paradigm—intelligence, critical thinking, creativity, and empathy—with machine strengths that include processing power and speed, information volume, pattern recognition, root cause analysis, forecasting, and prescriptions across a full spectrum of complex data. Furthermore, the human-machine teaming paradigm mitigates many human cognition weaknesses and shortfalls, as machines understand data as a series of discreet events instead of through stories and
storytelling, and do not actively search for preconceived patterns in randomness that influences human decision-making. In fact, as renowned psychologist Daniel Kahneman concluded in *Thinking, Fast and Slow*, only a considerable investment in time can possibly offset human cognitive limitations, which presently is limited to personal education, awareness, and practice. Due to greater machine vision across vast amounts of structured and unstructured data, this paradigm offers the opportunity to separate facts from human opinion when analyzing data, and better neutralize the human propensity to bound rationality. Therefore, this paradigm is an advanced approach that would evolve the human-machine relationship well beyond the current human use of machines to search for information, and instead to the human ability to better reason across data information with a machine.

Accordingly, the human-machine teaming paradigm indicates the ability to collectively elevate the rungs of Bloom’s Taxonomy for the cognitive domain, as superior strategic judgement would result with the combination of the superior soft skills of the human with the superior hard skills of the machine. This augmented cognition would enable humans to better focus on the “moral forces” that Carl von Clausewitz called “the most important of war”—including will, motivation, and creative genius. Additionally, machines will greatly enable more human “reflective openness,” as humans will have greater awareness of biases, limitations, risks, and underlying rationales in strategic thinking. This paradigm will also lead to better systems thinking, with the human-machine team examining the linkages and interactions between the components that comprise the entirety of the system. Therefore, the human-machine team will discover and reveal insights across large and complex data sets, elevating
critical thinking to inspire creative synthesis and evaluation of key factors to increase the probability of a desired outcome.

Just as modern human-machine teaming leads to superhuman performance in engineering and design, human-machine teeming will lead to the superior design of strategic ways, concepts, and decision-making. A recent example of the power of the human-machine teaming’s ability to collectively generate models, assess risk, and develop superior ways stems from what has become known as “Kasparov’s law.” In Deep Thinking, Garry Kasparov—the international chess champion defeated by IBM’s Deep Blue machine in 1997—examined the results from a 2005 freestyle chess tournament where anyone could compete in teams with other players or computers. Competitors included individuals and groups of chess Grandmasters armed with modern computer capabilities, beginners with less advanced technological capabilities, and even individual machines. And as Kasparov highlighted:

The winner was…a pair of amateur American players…using three computers at the same time. Their skill at manipulating and “coaching” their computers…counteracted the superior chess understanding of their Grandmaster opponents and the greater computational power of other participants. It was a triumph of process. A clever process beat superior knowledge and superior technology. Kasparov would conclude with an abstract formula—now referred to as Kasparov’s law—that a “weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process.” Thus, the victory was not due to a division of labor of chess pieces or responsibilities between player and machine, but instead because of the perpetual collaboration between the player and machine to yield a superior process.
Given this example, one can envision how human-machine teaming will augment human cognitive abilities to generate superhuman processes—or ways—that can best even the most superior knowledge and superior technology. Therefore, the role of the human-machine paradigm of strategy decision-making is generating concepts across the levels of war (see Figure 4). At the strategic level, human-machine teaming will yield a more complete understanding of the strategic environment through augmented systems thinking. Empowered by super-human critical, creative, and systems thinking at the strategic level, the human-machine team will then generate strategic forecasts, appraisals, and scenarios while producing strategic joint concepts and options. At the operational level, human-machine teaming will enable mission analysis, problem root cause analysis, and center of gravity analysis to generate multi-domain concepts, courses of action, and approaches. Moreover, human-machine teaming will enable a more accurate characterization and assessment of risk across the levels of war to include suitability, acceptability, and feasibility.

**Machine-Only Paradigm: Employment**

While one can envision a machine-only paradigm for strategy decision making, the paradigm is not appropriate for consideration at the strategic and operational levels. As highlighted earlier, strategy decision-making has a vital social complement. Without a social element to AI—such as understanding social intelligence concepts including
politics, relationships, or empathy—a machine-only paradigm would feature an algorithm or model that is unable to make decisions on behalf of humans regarding national-level strategy. Furthermore, machines will continue to lack the critical skill of creativity that is an essential element to strategy decision-making. While AI has mastered the ability to identify complex patterns that are beyond the ability of human senses, the simple fact is that “AI still cannot master everyday creative skills.” Most importantly, even if AI developed a social intelligence and creative ability—which would signal a coming technological singularity—it is ethically unimaginable to even consider AI making decisions at the strategic and operational levels without humans.

However, the machine-only paradigm is more appropriate to replace humans with automation of very specific objectives, tasks, or scopes found at the tactical level. In June of 2016 an AI-based pilot running on a $35 Raspberry Pi defeated a retired U.S. Air Force fighter pilot in a simulated air-to-air dogfight, so the question is not a matter of when AI will replace specific military objectives, tasks, or scopes, but instead what humans will permit AI to replace. Small, scalable, swarming autonomous air and space systems will one day conduct air and maritime superiority, global strike, and amphibious assault missions. Autonomous cybersecurity systems will rapidly identify, block, or destroy malicious attacks. Autonomous ground systems will one day storm an enemy position, and rapidly coordinate fires and maneuver via networked platforms.

Therefore, the machine-only paradigm is best reserved for tactical-level employment strategies (Figure 5) such as highly coordinated, multi-domain system battles against physical or logical targets. Department of Defense directives presently restrict the development and employment of systems with certain autonomous
capabilities to ensure a human is “in-the-loop” regarding decisions for all uses of lethal force. Accordingly, the human-only paradigm will continue to hold a monopoly on multi-domain command at the tactical level. Offensive autonomous systems—with a human in-the-loop—will engage in multi-domain engagements and battles. The remaining tactical actions will strike the right balance between the human-machine teaming (i.e., the “human in the loop”) and machine-only paradigm of autonomy. Thus, autonomous, tactical-level systems will rapidly synthesize data to understand the tactical environment, control the movement and maneuver of forces to gain decisive advantages, and conduct logistical sustainment of fielded forces.

"iStrategy: Judgement Superiority through Human-Machine Teaming"

Air Force Chief of Staff General David L. Goldfein recently set an urgent strategic tone by stating that “we’re transitioning from wars of attrition to wars of cognition,” and “our job is actually not to think outside the box, but to throw away the box.” The human-machine teaming paradigm indicates that AI will significantly augment human strategy decision-making abilities and yield a new source of national power—cognitive power. Human-machine teaming also represents the opportunity to significantly evolve human relationships with computers from search to reason with better use of data and information, and therefore transition from present-day information superiority to next generation judgment superiority. The pinnacle of judgment superiority would be human-machine cognitive power that leads to the generation of “iStrategies” to complement existing Department of Defense offset concepts and proposals.
While information and data serve as the fuel of judgment superiority through human-machine teaming, presently the Department of Defense narrowly views data and information as an asset to be made available and accessed, and therefore not from the context of how information could be translated into cognitive power. For example, information superiority is defined in Joint Publication 6-0 as “the operational advantage derived from the ability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary’s ability to do the same.”

Moreover, the fact that this “operational advantage” falls under a communications systems support publication also demonstrates the impression that management of information is the priority—and that the concept of information superiority is presently more aligned to wars of attrition than wars of cognition.

Furthermore, the Department of Defense’s new offset strategies are primarily focused on strategic means and lack any strategic-level proposals for generating advanced ways and concepts. The third offset strategy is the most recent proposal to invest in innovation to counteract and offset adversarial anti-access and area denial (A2/AD) advantages in the operational environment. The third offset strategy—combined with the Defense Innovation Initiative designed to leverage commercial advances in technology—is designed to increase the competitive advantage of US forces in the coming decades by exploiting unmanned operations, low-observable and extended air operations, undersea warfare, and the operational integration of resources and capabilities. Therefore, the third offset strategy leverages tactical-level automated and autonomous systems—and thus the means of strategy—which is more appropriate for the machine-only paradigm.
The human-machine teaming paradigm for strategy decision-making would lead to a competitive advantage in the development of iStrategies that would be designed to bolster conventional deterrence, defeat, and assurance by focusing on the creation of superior ways and concepts to apply military power. The third offset strategic goals focus on deterrence through denial and subsequent punishment, with an additional emphasis on assurance of allied partners. The primary target of third offset deterrence is the adversary’s perception by either decreasing will (denial), or increasing the enemy costs of potential attacks (punishment). Success in deterrence and war is directly linked to influencing an adversary’s will, for as characterized by Clausewitz, a nation’s power to wage war or resist is equal to total means available times the strength of will. By leveraging insights across the boundless amount of complex data sets and information, the human-machine teaming paradigm would generate iStrategy ways that can beat even the most superior knowledge and technology by an adversary as characterized by Kasparov’s Law. Furthermore, a Department of Defense competitive advantage in human-machine teaming for strategic decision-making—and thus judgment superiority—will influence and shape the perceptions and will of not only adversaries, but also allies. Therefore, transitioning to judgment superiority through human-machine teaming represents the best way to “throw away the war of attrition box” and augment the next generation iStrategists to shape and prevail in wars of cognition.

Challenges for the AI-Enabled Paradigms

There are many challenges ahead to realize the AI-enabled paradigms, which starts with negative perceptions of AI, moral questions, and organizational culture concerns. AI has historically suffered from unrealistic hype and expectations, which has generated public pessimism, anxiety, and even fear regarding the ability for AI to
become sentient and potentially even harm humans. AI-enabled paradigms in strategy decision-making also raises questions regarding traditional definitions of morality and ethics, from ensuring AI conforms to existing domestic and international laws and norms, to protecting the safety, privacy, and security of citizens. Additionally, entrenched military service cultures—with well-fortified human-centric approaches that favor traditional capabilities such as piloted aircraft, tanks, and aircraft carriers—have so far been extremely reluctant to embrace autonomy and autonomous systems, much less a concept of AI enabling strategy decision-making.

Poorly stated or ill-defined political goals, moral constraints, and risk tolerances could also lead to AI generating narrowly specified ways that leads to unanticipated or even unacceptable consequences. Strategic leaders will be under pressure to not only represent the military perspective candidly and apolitically, but must also significantly engage senior leadership when strategic guidance is too broad for subsequent translation into the criteria required to realize any strategic-level benefits from AI. Additionally, what strategically “matters”—such as the adversary’s will—would be very difficult to translate with a machine. For example, a machine would have challenges with understanding human social concepts such as measuring an adversary’s honor or resolve, or the ability of a nation to deter or coerce an adversary. Therefore, humans would need to leverage creativity abilities to translate and define abstract matters into criteria understood by machines.

Furthermore, AI-enabled paradigms will face significant challenges of integrating, verifying, and validating the data, information, and resulting models. While data and information are vast and broad across the Department of Defense, the data and
information is certainly not integrated and is instead “stovepiped” across organizational networks, drives, and databases. Next, while machines in theory are much less likely to bias information, the fact remains that machines are still programmed by humans. Even the most aware human programmer could inadvertently bias the model or results, giving a machine a new ability to further amplify human biases and ignorance. Additionally, AI is only as effective as the data from which it initially trains, so AI that learns from a biased data set will only lead to a biased capability. Most dangerously, the model could suffer from injections of “poisoned” or forged data inputs during training or execution. Finally, an AI could confront conditions in complex or uncertain environments that were never contemplated during initial training and testing, which could lead to unpredictable behavior in the model.

Finally, one of the biggest challenges to realizing the machine paradigms is the present shortfall of available AI expertise. Commercial industries including marketing, finance, transportation, aviation, telecommunications, manufacturing, and health care are now extensively leveraging the power of AI to transform a spectrum of services. Moreover, significant expertise is required for developing and supervising AI, and that talent is significantly limited commercially and especially in the government sector. Unfortunately, an ensuing “talent war” has subsequently erupted in the commercial sector since AI has now proven the ability to generate financial profits.

Recommendations and Conclusion

Recent U.S. national guidance directs the Department of Defense to prioritize investment in AI to gain competitive military advantages. Per the 2017 National Security Strategy, the U.S. will “lead in research, technology, invention, and innovation” by “prioritizing emerging technologies critical to economic growth and security, such
as...artificial intelligence.” The **National Security Strategy** also highlights “the field of artificial intelligence, in particular, is progressing rapidly.” Additionally, the **2018 National Defense Strategy** underscored that the changing character of war demands new technologies including “advanced computing, ‘big data’ analytics, artificial intelligence, autonomy, robotics” to “ensure we will be able to fight and win the wars of the future.” Finally, the **National Defense Strategy** also emphasized, “the Department will invest broadly in the military application of autonomy, artificial intelligence, and machine learning, including rapid application of breakthroughs, to gain competitive military advantages.”

However, the Department of Defense presently has a fragmented, decentralized posture that risks the strategic opportunity of leveraging emerging advances in AI to meet national-level guidance and intent. Secretary of Defense James Mattis has recently commented that while the Department has an extensive history of integrated technology, integration of AI has lagged the commercial sector. In fact, it is likely more accurate to state that the Department of Defense is generally AI illiterate except a few elements within defense directorates, service departments, agencies, and combatant commands that are pursuing independent, experimental approaches to AI. For example, the Assistant Secretary of Defense for Research and Engineering provides broad guidance on AI experimental investments across the defense research and engineering enterprise—including the Defense Advanced Research Projects Agency and service laboratories—to extend the capabilities of existing military systems and investigate new technological innovations. Additionally, the Under Secretary of Defense for Intelligence, Combat Support Agencies such as the National Security Agency, and the U.S. Air Force

19
Intelligence Directorate are all experimenting with AI to enhance intelligence collection, processing, exploitation, analysis, and dissemination.

While the Department of Defense lags behind the commercial sector regarding AI innovations and integration, a more dangerous perspective is the potential for the Department to lag behind global adversaries such as Russia and China. For example, Russian President Vladimir Putin recently set a noteworthy tone on AI by remarking that “the one who becomes the leader in this sphere will be the ruler of the world.”

In late 2017, China quietly released and began implementing its own *Next Generation Artificial Intelligence Development Plan* modeled after the significant amounts of research, lessons learned, and expert recommendations made by the National Science and Technical Council during the Obama administration.

Therefore, the Department of Defense now requires unity of effort on AI to posture the department for competitive military advantages—such as judgment superiority through cognitive power—and transition from primarily experimenting with AI to instead *pioneering* and *operationalizing* AI. This starts with creation of a new Deputy Assistant Secretary of Defense for AI (DASD/AI) to create, integrate, and align Department of Defense AI strategies, policies, resources, and programs. The DASD/AI would serve under the DASD for Strategy, Plans, and Capabilities in the Office of the Under Secretary of Defense for Policy, with roles and responsibilities similar to the DASD/Cyber (see Appendix 1 for a full list of draft DASD/AI roles and responsibilities).

The DASD/AI would integrate and synchronize Department of Defense stakeholders by developing an AI vision and strategy. Thus, using the creation of a first-ever *Department of Defense Strategy on AI* as the forcing function, the DASD/AI would
establish a formal coalition through an AI Executive Steering Group of senior stakeholders from department directorates, services, agencies, and combatant commands to integrate, synchronize, and align Department efforts. The purpose of the AI Executive Steering Group would be not only to design and align a Department vision and strategy for AI, but to also ensure the stakeholders engage in the dialogue required to align the spectrum of required AI concepts, policies, programs, and processes. For example, the AI Executive Steering Group would address strategy elements and alignments in the following table:

<table>
<thead>
<tr>
<th>Elements of the DoD Strategy on AI</th>
<th>Additional DASD/AI Alignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define a primary purpose, vision, and mission for defense AI</td>
<td>Envision broad concepts of how AI will enable and integrate with military operations and contingency plans</td>
</tr>
<tr>
<td>Direct broad priorities and goals for AI research and development</td>
<td>Develop broad awareness of operational requirements by integrating into JCIDS</td>
</tr>
<tr>
<td>Direct collaboration priorities and goals across defense agencies, laboratories, offices, and cross-functional teams involved in AI</td>
<td>Review existing defense AI research and development initiatives</td>
</tr>
<tr>
<td>Establish private sector integration priorities and goals including the technology, telecommunications, services, and defense industries</td>
<td>Characterize USG partnerships required for defense AI</td>
</tr>
<tr>
<td>Characterize anticipated defense AI joint forces, force postures, and force structures</td>
<td>Outline talent management priorities to recruit, train, and retain the professionals required for the development and supervise on of AI</td>
</tr>
<tr>
<td>Establish joint force training and development priorities and goals for AI</td>
<td>Establish a notional baseline for defense AI to fund, acquire, distribute, expend, and account for future financial resources</td>
</tr>
<tr>
<td>Establish modernization priorities and goals for AI system acquisition and procurement</td>
<td>Outline policies regarding morality and ethics, including how defense AI will conform to domestic law, and international laws agreements and legal regimes, while protecting the safety, privacy, and security of U.S. citizens</td>
</tr>
<tr>
<td>Establish information architecture priorities to better integrate public and private “big” data and information</td>
<td>Outline the general processes for developmental and operational testing,</td>
</tr>
</tbody>
</table>
The DASD/AI would use the final *Department of Defense Strategy on AI* to lead and communicate the change required across the Department through the establishment of subsequent goals and objectives to implement the strategy. The AI Executive Steering Group would therefore evolve to include a dialogue on the actions required of DoD stakeholders to realize the goals and objectives of the strategy.

Regarding AI, Secretary Mattis recently stated on 17 Feb 18 that “I’m certainly questioning my original premise that the fundamental nature of war will not change…you’ve got to question that now.” The military services will soon rely on AI to enable everything from national-level strategy decision-making to tactical-level engagements and battles. Therefore, the Department of Defense can no longer delay thinking strategically about future human relationships with computers and how to better use data and information. A critical first step is to establish a new DASD/AI to create, integrate, and align Department efforts in order to meet national-level guidance and intent on AI. Now is the right time to evolve organizationally to leverage advances in AI to extend human cognition and augment human decision-making capabilities to defend the U.S. in the coming AI age.
Appendix 1: Draft Roles and Responsibilities of DASD/Al

- Office of primary responsibility within OSD for policy matters related to AI activates or involving AI systems
- Ensure AI-related activities are integrated into national and DoD strategies
- Develop, coordinate, and oversee implementation of U.S. government and DoD policy and strategy for AI operational activities and systems
- Formulate specific DoD policies and guidance on:
  - AI forces, systems, and their employment
  - Integration of cyber capabilities into operations and contingency plans
  - International, U.S. government, and DoD AI operations cooperation
  - International agreements and legal regimes
- Review and evaluate AI-related programs, plans, and system requirements
- Participate in planning and budgeting activities for AI operations and systems
- Represent OSD at interagency deliberations and international negotiations
- Interface with other U.S. government departments and agencies, Congress, the public, and foreign governments on AI policy matters
Endnotes


15 Allen and Chan, Artificial Intelligence and National Security, 27.


17 Dean, “How Will Artificial Intelligence Affect Your Life.”


29 Ibid., 30.


34 Ibid., 61.


41 Clausewitz, *On War*, 579.


50 Andrew Hill, *The Devil You Know: Strategic Thinking in Complex Adaptive Systems*, Faculty Paper (Carlisle Barracks, PA: U.S. Army War College, August 2017), [https://armywarcollege.blackboard.com/bbcswebdav/institution/REP/Course%20Files/DCLM/Co](https://armywarcollege.blackboard.com/bbcswebdav/institution/REP/Course%20Files/DCLM/Co)

52 Kasparov, Deep Thinking, 246.

53 Ibid., 246.

54 Ibid., 246.

55 Sawyer, Explaining Creativity, 3.

56 Allen and Chan, Artificial Intelligence and National Security, 21.


58 Allen and Chan, Artificial Intelligence and National Security, 13.


64 Ibid., 39.
65 Clausewitz, *On War*, 77.

66 NSTC, *The National Artificial Intelligence Research and Development Strategic Plan*, 27.

67 Gill, “Artificial Super Intelligence,” 139.


79 Ibid., 20.


81 Ibid., 7.


85 Ransbotham, Kiron, Gerbert, and Reeves, “Reshaping Business with Artificial Intelligence,” 6.