Issue Brief

AI for Military Decision-Making

Harnessing the Advantages and Avoiding the Risks

Authors
Emelia S. Probasco
Helen Toner
Matthew Burtell
Tim G. J. Rudner

CSET CENTER *for* SECURITY *and* EMERGING TECHNOLOGY

April 2025

Executive Summary

The integration of artificial intelligence into military operations has become a significant focus for armed forces globally. Military commanders are interested in AI's potential to improve decision-making, especially at the operational level of war, where they must integrate a lot of information quickly to make life-and-death decisions. However, the enthusiasm for AI-enabled decision support systems (DSS) must be balanced with an understanding of their capabilities and limitations to ensure appropriate and effective deployment. This paper reviews recently proposed uses of AI-enabled DSS, provides a simplified framework for considering AI-DSS capabilities and limitations, and recommends practical risk mitigations that commanders might employ when operating with an AI-enabled DSS. Our framework for considering AI-DSS intended for operational decision-making emphasizes three critical areas.

1. Scope considerations: Is the scope of the AI-DSS well-defined and understood?

- **Context shifts.** AI systems are prone to fail if used in settings that are meaningfully different from their training data.
- **Projection and prediction.** There is an important difference between predictions based on physical laws and those involving human interactions. For the latter, we lack accurate models and directly observed data.
- **Flexible or unclearly scoped systems.** AI-DSS can be flexible, but without well-defined use cases or guardrails they can confuse operators and lead to misuse.
- **Irreducible uncertainty.** Questions like "What will the enemy do?" have an inherent level of uncertainty that cannot be fully eliminated. DSS users must understand that certainty is an impossible goal and human judgment is still required when using AI-DSS.

2. Data considerations: Does the training data substantiate the AI-DSS' conclusions?

- **Quality and fidelity.** High-quality, relevant data is critical for effective AI systems but challenging to gather and maintain. Human behavior data is particularly challenging to use effectively, due to its indirect observability and demographic variability.

- **Skewed data.** Military commanders struggle to obtain accurate data on friendly and enemy forces. Data biases—such as those arising from sensor availability, deception, or in social media—can significantly impact AI system outputs.

- **Scarce data.** The ability of an AI system to provide analysis or predict outcomes in combat or war may be limited because data about combat and war is limited. Traditional intelligence methods, which can combine insights and inferences that rely on a richer understanding of the relevant context, can be more valuable.

3. Human-machine interaction: What are the capabilities and limitations of the human-machine team as a single system within a given context?

- **False expectations with LLMs.** Large language models (LLMs) are powerful tools, but they must be applied with care as they can mislead users by confidently presenting incorrect information, fabricating justifications, or increasing user acceptance of erroneous recommendations.

- **Human biases.** Users must understand how their cognitive biases—such as automation bias, confirmation bias, or recency bias—may be affected by AI-DSS outputs, especially in stressful scenarios.

- **Organizational biases.** Overreliance on DSS due to perceived ease of use can lead to poor decision-making, especially in extreme situations when risk tolerance is high. Organizations must be careful to avoid hasty or under-resourced decision-making based on a false perception of AI capabilities.

Based on our analysis, we recommend the following risk mitigation strategies when using AI-DSS:

1. Set context- and risk-based criteria for deployment: Commanders should set the time, place, and context for authorizing DSS use and prepare forces to adapt software settings as conditions and risk tolerance change.

2. Train and qualify AI-enabled DSS operators: Operators should be thoroughly trained on DSS capabilities and limitations. Those involved in lethal operations should undergo examinations for official qualifications appropriate to their role in operating the system.

3. Establish a continuous certification cycle: Units leveraging AI-DSS should be regularly certified to reduce the risk of inappropriately deploying or operating the system. Sharing performance metrics with data scientists, operations analysts, and experts in continuous tests and evaluations can help validate continued responsible use of AI-DSS and also support technical evolutions.

4. Designate a Responsible AI officer: Akin to establishing safety and mishap programs, Responsible AI (RAI) officers in military units can serve as local conduits for new information, promoting broad-based AI literacy, reporting AI incidents or mishaps to a higher authority, and mitigating AI-DSS risks.

5. Document incidents and harms: Documenting AI system flaws and user mistakes is essential for avoiding repeat errors and for building trust through transparency. RAI officers should be responsible for such documentation, akin to mishap reporting processes already in place in the services.

The integration of AI into military decision-making presents both opportunities and challenges. By carefully considering the scope, data quality, and human-machine interaction, and by implementing rigorous training, certification, and safety measures, military organizations can leverage AI more effectively while mitigating potential risks.

Table of Contents

Introduction

Artificial intelligence promises to help military commanders make sense of vast amounts of data at superhuman speeds. The military has a strong motivation to take advantage of AI, and among the most interested within the armed forces are commanders charged with making operational decisions in war. These commanders must continuously "observe, orient, decide, and act" on a fast-paced and multidimensional battlefield where decisions are life-and-death.

The historical desire for sophisticated tools to maintain battlefield awareness, support military planning, and even predict future enemy movements or reactions has led to the creation of everything from weather modeling to campaign modeling to early warning systems.[*]

The strong and understandable desire for AI-enabled decision support systems (DSS) must be tempered, however, by an understanding of the capabilities and limitations of these systems, which should dictate when and how they are deployed.

We begin with a brief history of efforts to fight through the fog of war and the emergence of decision frameworks with supporting tools, linking these to the recent quest for AI-DSS. We then demonstrate the widespread interest in applying AI for decision support among the world's most powerful militaries. Finally, we characterize the opportunities and risks of applying AI to military decisions and offer a basic framework to guide the deployment of these systems.

[*] The Royal Navy served a leading role in the establishment and use of weather forecasting, including the luminaries Sir Francis Beaufort (for whom the Beaufort scale for wind force is named) and Vice Admiral Robert FitzRoy, who founded the Royal Meteorological Society in the 1850s.

Background and Historical Efforts toward Awareness and Prediction for Military Decisions

Military commanders have sought support for their decisions since ancient times. Even Herodotus's *Histories* describe how Croesus consulted the oracles of Greece and Libya in the sixth century BCE when he decided to go to war. When he asked "if he should undertake an expedition against the Persians," the oracle assured him that "he would destroy a great empire."[1] While Croesus successfully destroyed a great empire, it turned out to be his own.

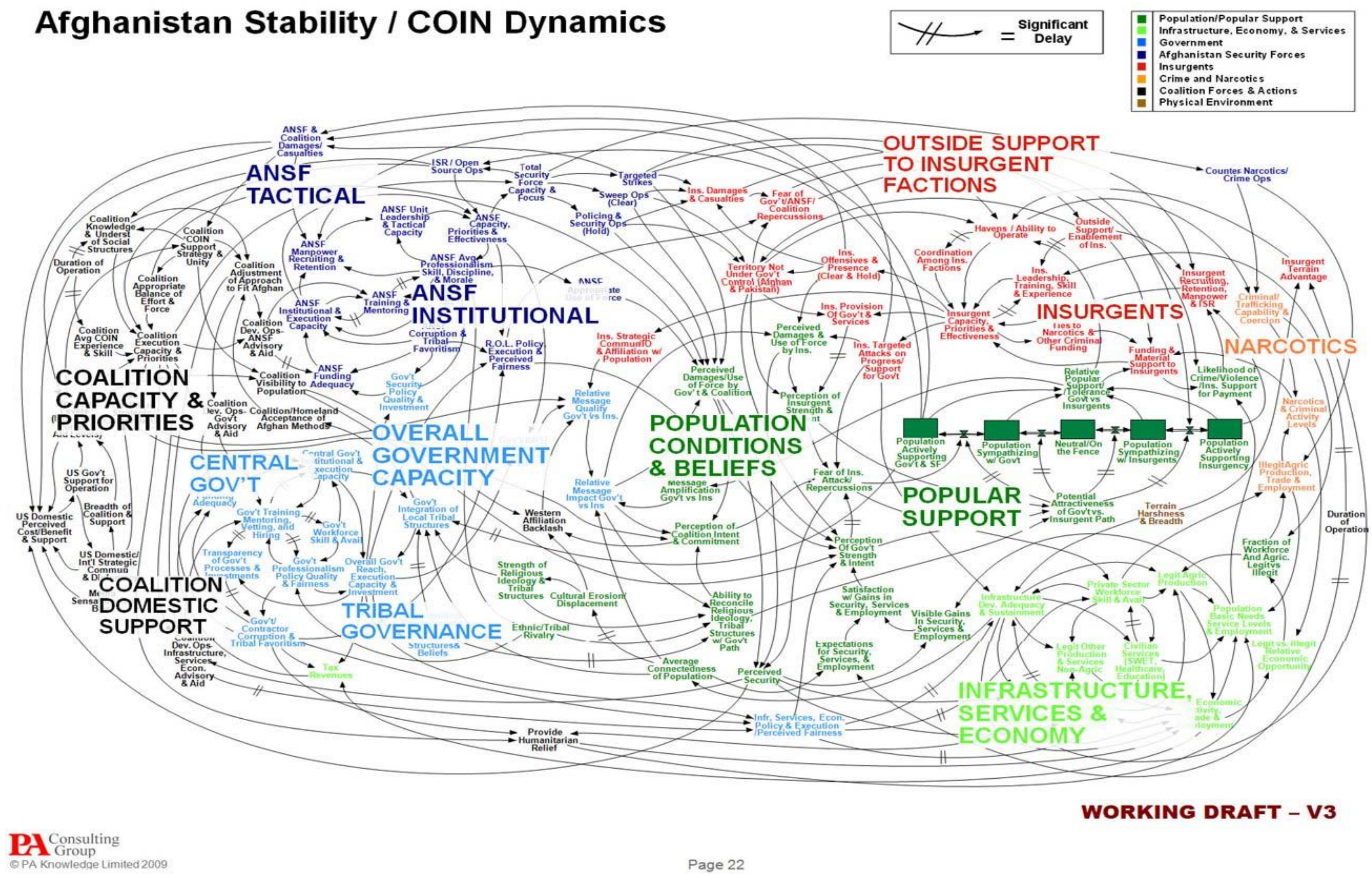
While oracles and AI-DSS are very different, both have been consulted by commanders who must make high-stakes battlefield decisions with imperfect information. Today, thanks to the proliferation of data sources and sensors, the information is both overwhelming and imperfect. Commanders seek to understand the quantities, capabilities, operating status, locations, and logistics of their own forces, while gathering intelligence on the enemy and being careful to avoid civilians.

This information burden—the need to constantly gather data and integrate it to make timely decisions—has long been known as the "fog of war," because it is simultaneously incomplete, confusing, complex, and continually changing.

Evidence of efforts to collect and systematically process information relevant to military decisions can be found in doctrinal memos, training books, and PowerPoints from recent conflicts. These include everything from the famously complex chart once briefed to the commander of U.S. forces in Afghanistan (Figure 1) to rigidly simplified approaches in Russian textbooks (Figure 2). In these and other examples, it is clear that a wide variety of military, economic, social, and political intelligence sources must inform a commander's choices in a war.

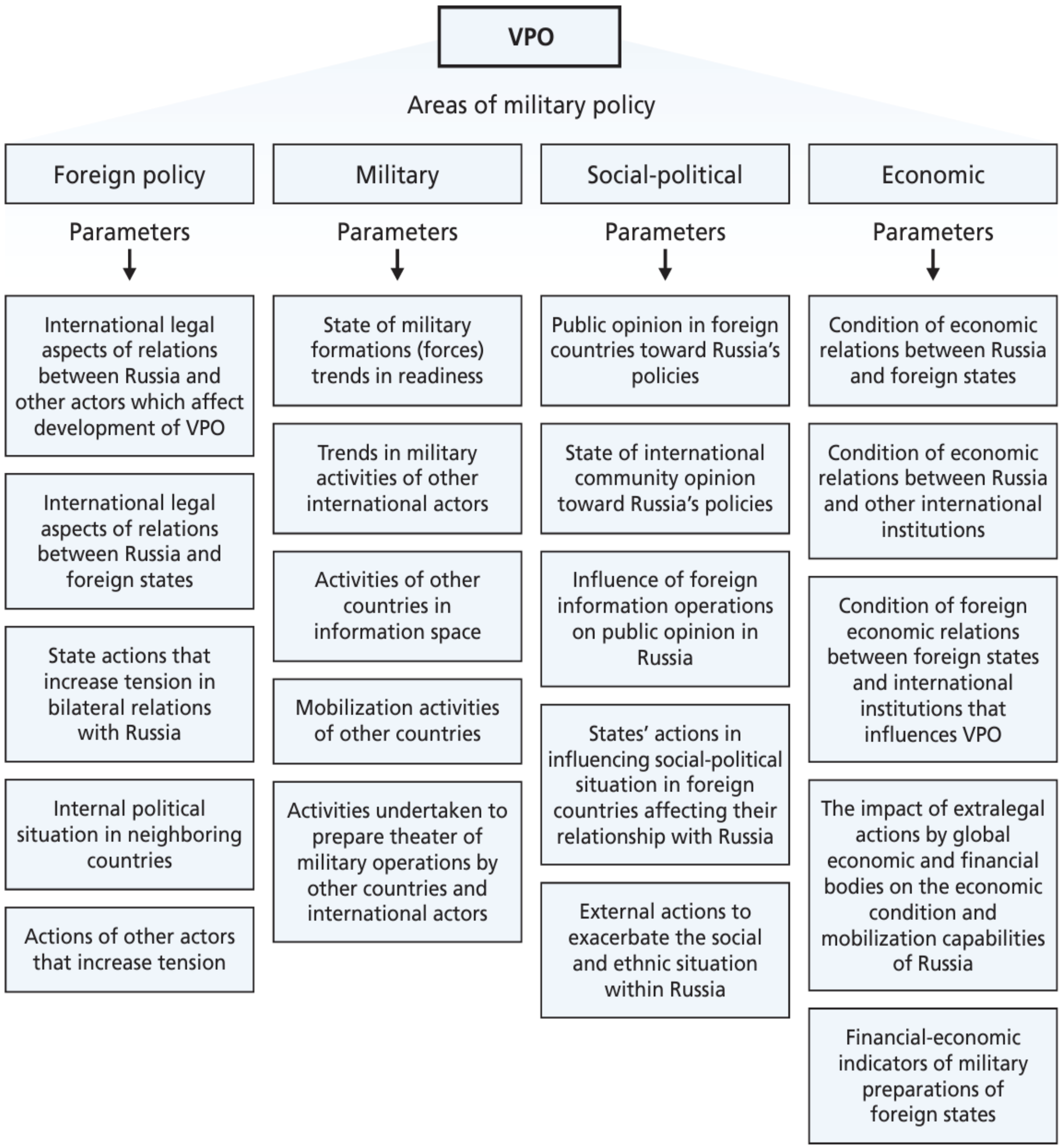
This paper is specifically focused on the operational-level decisions military commanders must make, though the difference between tactical, operational, and strategic military decisions is not always clear. In general, the operational level of war addresses higher-level objectives—such as the orchestration of logistics, target generation for artillery fires, or coordination of multiple units—without rising to the strategic level, which considers campaigns and major operations on the national scale.[2] Examples of DSS we will not address include tools intended to guide acquisition and investment decisions, to support situational awareness in peacetime, or to select tactical targets (i.e., an assault rifle advanced guidance system like ARCAS).[3]

Figure 1. PowerPoint Slide as Presented to International Security Assistance Force Commander



Source: Amy Davidson Sorkin, "Close Look: Mapping the War," *New Yorker*, December 10, 2009.

Figure 2. Russian Textbook Approach



Source: Clint Reach et al., "Russian Military Forecasting and Analysis," RAND, June 23, 2022.

The operations research and engineering communities have developed mathematical models in the past to help address components of operational decisions for commanders.[4] Some previously developed models are narrowly focused and physics based, such as those aimed at predicting the outcome of aerial dogfights, the effectiveness of a certain weapons against a target (weapon-target pairing), or how ordinance might explode and create damage (fragmentation prediction). Other more complex models, such as campaign models, seek to support planning and predict campaign-level outcomes. These combine multiple data sources and models that are not solely based on physical limitations but include statistical inferences based on historical data and/or assumptions.[5] These include IDAGAM, TACWAR, THUNDER, JICM, and STORM (all campaign models known by their acronyms).[6]

Research has shown that few non-physics-based tactical or operational models have ever been adequately validated.[7] Despite this obvious shortcoming, imperfect models, from campaign analysis to war-gaming, are still used. While overreliance on these unvalidated models is dangerous, they are seen as helpful for structuring decisions and challenging decision maker assumptions.

Global Efforts to Adopt AI for Military Decision Support

Recently, advancements in sensors, data sources, and algorithms have revived hopes for a tool that can accurately assess the battlefield and predict likely events or campaign outcomes. While AI and information technology will bring significant capabilities to support operational commanders, machines still cannot perfectly understand or predict war.[8] Making clear distinctions between what AI can and cannot deliver to military decisions is imperative to their effective deployment.

Statements from militaries worldwide are indicative of the high (and sometimes unrealistic) expectations of what AI-DSS can do (see Table 1 for an illustrative list). Beyond mere statements, governments have invested in developing AI-DSS. Russia already relies on "algorithms" to assess forces and determine courses of action in a highly formulaic way.[9] China's People's Liberation Army has issued calls for proposals for "battalion and company command decision-making model and human-machine teaming software."[10] The North Atlantic Treaty Organization (NATO) Science and Technology Office has initiated a research team to "investigate how reinforcement learning could be used to support decision-making."[11] The Defense Advanced Research Projects Agency (DARPA) has its "In the Moment" program, which aims to help human decision-makers in medical triage emergencies.[12] These are just a few examples of the range of applications and global nature of development.

Table 1. Illustrative Statements about DSS from Global Powers

Australia	"Providing high-level strategy decisions is currently beyond the state-of-the-art of AI, but there is ongoing work to broaden the applicability of AI techniques to support the humans making these complex decisions. We believe the ADF [Australian Defence Force] would benefit from following these developments closely and investing as appropriate. It is likely that adversaries who embrace such technology will have a dramatically reduced decision-making cycle as the capabilities in this area improve."[13]
China	"Build new generation AI based on research and development in the common theory and critical common technology. Establish mechanisms to normalize communication and coordination among scientific research institutes, universities, enterprises and military industry units. Promote military-civilian two-way transformation of AI technology. Strengthen a new generation of AI technology as a strong support to command and decision-making, military deduction, defense equipment, and other applications. Guide defense domain AI technology toward civilian applications."[14]
Russia	"AI technologies can change perceptions about the size of anticipated costs and expected benefits, the balance between offensive and defensive measures, and the results of conventional and nuclear deterrence calculations, eliminating uncertainty in situation assessment, ensuring near-absolute impartiality in political and military decisions, and completely eliminating the influence of the human factor."[15]
NATO	"Improving the Alliance's situational awareness and strategic anticipation has been an important dimension of the Alliance's strengthened deterrence and defence posture. Fundamental to the Alliance's ability to shape, contest and fight is expanding knowledge and understanding, with a view to ultimately achieving cognitive superiority. This understanding shall be connected across all-domains, enabled by technology, in order to maximize commanders' ability to anticipate, think, decide and act. Efforts to build better situational awareness and understanding with a view to achieving cognitive advantage over potential adversaries is a priority for the Alliance."[16]
United States	"The latest advancements in data, analytics, and artificial intelligence (AI) technologies enable leaders to make better decisions faster, from the boardroom to the battlefield. Therefore, accelerating the adoption of these technologies presents an unprecedented opportunity to equip leaders at all levels of the Department with the data they need, and harness the full potential of the decision-making power of our people."[17]

Globally, companies have responded to these government statements of interest, and there is now varied evidence of DSS being marketed or employed. In the United States, Palantir markets AIP as a system that integrates disparate data streams and applies machine learning to support commander situational awareness and target identification. Also in the United States, Scale AI's Donovan is marketed as a chatbot that can be fine-tuned on sensitive or classified government data. The French multinational Thales Defense offers the AI-enabled decision support tool ANTICIPE, which integrates a "wargaming tool and advanced machine learning algorithms to provide military commanders with actionable insights when they need it most."[18] Finally, China-based StarSee's "Real-time Combat Intelligence Guidance System" offers to integrate intelligence to support commander decisions.[19] These are just a few examples of global DSS. No two appear exactly the same, but all endeavor to make sense of the battlefield and support better decisions.

Types of Decision Support

The range of AI-DSS being advertised today blurs the lines between different decision support tasks, from awareness to planning to prediction, as well as the lines between operational- and strategic-level decisions.[20] Examples include systems that use computer vision and data fusion to identify friendly and enemy forces on a battlefield, algorithms to spot anomalies, algorithms to perform sentiment analysis on social media text, and, more recently, large language models (LLMs) for evaluating intelligence or generating courses of action (COAs).

Marketing videos and military leaders have discussed applying these tools to artillery targeting, route planning, logistics optimization, and the apportionment of medical care in emergencies. Some proposals have even postulated that AI might be used to predict the likelihood of success of a chosen COA or the likelihood of a future conflict or political instability. Based on a survey of largely American AI-enabled DSS (the Appendix contains a sample), Table 2 provides an illustrative list of tasks that have been proposed by government and industry leaders, some of which are already available. Notably, some systems commingle operational- and strategic-level decision support, as illustrated in the list.

Table 2. Illustrative List of Advertised AI Applications for Military Decisions

- Situational awareness
  - Mapping friendly and enemy forces and maneuvers
  - Identifying anomalies in financial transactions or ship movements
  - Surfacing investments or academic research in novel technologies
  - Mapping public sentiment and summarizing news media coverage
  - Mapping and monitoring supply chains, financial flows, and social networks
- Planning and execution
  - Red-teaming analytical findings
  - Planning theater resupply and resiliency
  - Mass casualty responses
  - Fires/artillery target selection and weapon-target pairing
  - Generating and red-teaming COAs
  - Command and control of remote units
- Prediction
  - War-gaming
  - Predicting crises
  - Forecasting political instability, popular uprisings, and migration
  - Forecasting combat outcomes or COA success probabilities

While the AI tasks listed above are neatly categorized as awareness, planning and execution, and prediction, AI-DSS advertised by software developers often cross categories for both operational and strategic decisions. For example, Scale AI, Palantir, Anduril, and Rebellion Defense each promote comparable AI-enabled situational awareness tools that also support military planning or predict outcomes. These DSS leverage a number of AI techniques, including data fusion, computer vision (segmentation and classification systems), machine learning (anomaly detection and recommendation systems), and, most recently, generative AI and LLMs.

Besides these expansive systems that do several types of tasks, there is also evidence of much more narrowly scoped systems that may only address a few of the tasks listed in Table 2. For example, the 18th Airborne has used the Maven Smart System to support the artillery firing process, and the Intelligence Advanced Research Projects Activity (IARPA) runs the Rapid Explanation, Analysis, and Sourcing Online program, which is intended to evaluate and challenge the conclusions of intelligence analysts to improve their findings and arguments.[21] Researchers at the Naval Research Laboratory have advertised commercial licenses for a recommender system to help military decision-makers without data analysis skills better understand and analyze data.[22] And the company DEFCON AI markets a platform for logisticians to model supply chain disruptions and evaluate logistical resupply options (see the Appendix for more information).

These examples illustrate the types of efforts being proposed or marketed; we do not have exact information about the inner workings of each of these systems. Military decision-makers considering deploying DSS may face a dilemma if they also do not have access to—or do not understand—the details of how the systems work. Despite AI's many capabilities, commanders must carefully distinguish between use cases for which DSS are appropriate and those deserving more skepticism. To help address this challenge, we offer a framework for commanders to consider when evaluating the appropriateness and risk of employing AI for operational decision-making.

A Simplified Framework for Commanders Evaluating AI-Enabled Military DSS

A military leader who encounters AI-DSS for the first time may reasonably be excited by the clarity that such systems seem to offer and the relative speed and ease with which they can integrate and process vast amounts of data. These systems can indeed add clarity and support difficult decisions. Knowing when and where they are best positioned, as well as how to mitigate their inherent risks, is key to a military commander's responsible and effective use of them.

The U.S. Department of Defense has established lengthy guidance for military forces to help determine the responsible application of AI.[23] The guidance considers many facets of AI deployment, including consideration of available compute and specific AI tools and methods, for example. Similarly, past reports have gone into important details about the nuances of AI-DSS and their relationship to exercising human judgment in accordance with international humanitarian law.[24] While detailed guidance and nuance are important, this paper aims to serve as a simplified resource for decision-makers thinking at a high level about using or deploying AI-based DSS. The following three areas of inquiry are offered as a starting point.

1. **Scope**: How well-defined and understood is the scope of the system? What are the appropriate and inappropriate use cases?

2. **Data**: Does the training data substantiate the system's conclusions? How is the data collected and how frequently is it updated? What abnormalities, outliers, or irregularities might be present in the data?

3. **Human-machine interaction:** How does the human operator understand machine outputs? What are the capabilities and limitations of the human-machine team as a single system within a given context?

To dive deeper, we consider each question in turn and then recommend mitigations where appropriate.

Scope

How well-defined and understood is the scope of the system? What are the appropriate and inappropriate use cases?

Some of the AI-DSS available today are very tightly scoped, such as a tool that can use an image to identify sniper blind spots or one that reviews documents for foreign disclosure to allies.[25] The clearer and more well-defined the scope of an AI-DSS, the more confident a commander can be in applying the tool to the correct situation and avoid using it in ways for which it was not tested and validated. When considering whether the scope of a given AI-DSS is appropriate, commanders should be on the lookout for the following elements.

Context shifts. A general challenge with using modern DSS in high-stakes settings is that systems based on deep learning models—the primary AI paradigm in use today—are prone to fail if applied to settings that are meaningfully different from the settings they were trained on. For example, models previously successful at predicting shopping trends, traffic, and supply chains began to fail in 2020 as the COVID-19 pandemic spread and habits changed.[26] Similarly, it would be fraught to rely on sentiment analysis algorithms optimized for only one dialect of Arabic when trying to assess social movements in the Middle East.[27] In some instances, human operators can easily observe degradations in AI performance, but this is not always the case. Figure 3 depicts an example of a potential "distribution shift," drawn from research on identifying flooded buildings from overhead imagery.[28] A model trained on data from Houston would not necessarily perform well on data from Russia or Germany, where the landscapes in question are very different. Systems should be carefully tested under different operational conditions to ensure that the scope of situations where they perform well versus poorly is known and communicated to users.

Systems should be carefully tested under different operational conditions to ensure that the scope of situations where they perform well versus poorly is known and communicated to users.

Figure 3: Distribution Shifts in Image Processing: Overhead Images of Houston (United States), Orenburg (Russia), and Liers (Germany)

Source: Tim G. J. Rudner et al., "Multi³Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery," *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 1 (2019), https://doi.org/10.1609/aaai.v33i01.3301702; RFE/RL and Reuters, "Satellite Imagery Reveals The Raging Floodwaters Inundating Russia," Radio Free Europe, April 10, 2024, www.rferl.org/a/satellite-images-flooding-russia/32898717.html; Maxar Technologies, "Open Data Response to Flooding in Europe," Maxar (blog), July 21, 2021, https://blog.maxar.com/for-a-better-world/2021/open-data-response-to-flooding-in-europe.

Projection and prediction. Some decision support tools propose to ingest large datasets to predict the future, such as the possibility of social uprisings or military attacks. On the one hand, these sorts of predictions are not new: humans have long forecasted the weather, and today we can deliver highly accurate near-term predictions. But weather is a phenomenon that has been long observed, is well instrumented, and follows reasonably well-understood physical laws—and weather is still reliably predictable only a week or so ahead of time. Human interactions and decisions are far more complex, less understood, and not fully observable (see next section on data). AI-DSS whose scope includes making projections or predictions should be subject to additional scrutiny unless those predictions are based on well-understood physical laws and anchored in directly applicable data (e.g., predicting impact points based on ballistic missile trajectories).

DSS whose scope includes making projections or predictions should be subject to additional scrutiny unless those predictions are based on well-understood physical laws and anchored in directly applicable data.

Flexible or unclearly scoped systems. Having the ability to iterate software solutions for different contexts is a great strength, but flexibility can also be a liability for AI-DSS. This could be especially true when tactical or operational decision-makers try to leverage strategic-level DSS or, vice versa, when strategic decision-makers mistakenly use tactical or operational DSS. For example, models built for strategic-level investment decisions can rest on myriad assumptions or expert understanding that render them inappropriate for time-critical operational decisions or use by nonexperts. Similarly, a tactical DSS that leverages an LLM to summarize reports for a trained intelligence officer could not be reasonably or responsibly given to a senior operational commander who lacks the contextual knowledge needed to spot missed indicators, errors, and hallucinations.

LLM-based chatbots—with their natural-language interactions—can seem so universally capable (even when they are not) that their deployment should be accompanied by clear use guidelines and software guardrails.

Speaking more specifically about LLMs' issues of scope and flexibility, there is nothing inherently inappropriate about applying LLMs to decision support tasks. However, LLM-based chatbots—with their natural-language interactions—can seem so universally capable (even when they are not) that their deployment should be accompanied by clear use guidelines and software guardrails. Some early demos of

LLM-based products, in contrast, purport to be akin to an all-purpose "battle buddy" that could assist a single user with an exhaustive list of tasks, including identifying enemy units, searching the literature for contextual facts about combatant countries, analyzing intelligence, generating COAs, and directly sending digital commands to an autonomous or remotely piloted vehicle.[29] The breadth of these tasks, combined with the natural-language interface that encourages the operator to think of the system as akin to a human companion, is cause for concern. Without clear boundaries around which uses are appropriate, operators are likely to unknowingly stretch the limits of such systems, pushing them to assist with tasks that go beyond the contexts and purposes for which they were developed and validated.

Irreducible uncertainty. The Russian statement in Table 1 refers to the possibility of "eliminating uncertainty." This phrase demonstrates a belief, not limited to Russia, that it is possible, in principle, to reduce uncertainty to zero. In most real-world settings, this is not possible. Consider, for example, a fair coin flip. You could flip a coin repeatedly and—with sufficiently many coin tosses—would be able to become highly certain that the probability of a flip landing heads is 50 percent. However, this would not help reduce any uncertainty about whether the next flip will land heads or tails. In other words, the uncertainty is irreducible. The users of AI-DSS may want to know things like "What will my enemy do next?" or the real-world equivalent of "What move will my opponent play in rock, paper, scissors?" Questions like these have an inherent level of uncertainty that cannot be fully reduced.

The users of AI-DSS may want to know things like "What will my enemy do next?" or the real-world equivalent of "What move will my opponent play in rock, paper, scissors?" Questions like these have an inherent level of uncertainty that cannot be fully reduced.

Beyond the simplicity of coin flips, military operational decisions are complicated by the near impossibility of taking perfect measurements, modeling military operations, and accounting for nonbinary situations (e.g., What if the rock is pebble-size and the scissors are large and made of titanium?). In sum, this means that humans must still exercise judgment in making battlefield decisions. AI-DSS may reduce the number of unknowns and a degree of uncertainty, but commanders must also accept the limitations of these systems and understand that they cannot eliminate uncertainty.

Data

Does the training data substantiate the system's conclusions? How is the data collected and how frequently is it updated? What abnormalities, outliers, or irregularities might be present in the data?

Quality and fidelity. Modern machine learning systems for computer vision and natural-language processing tasks are developed using large quantities of data. When high-quality, relevant data is plentiful, building a system that works well is easier. Weather prediction is an example. We have enormous quantities of high-fidelity data on exactly this kind of phenomenon—temperature, air pressure, cloud formation, precipitation—and a strong mechanistic understanding of the physical dynamics that drive weather patterns. Importantly, the data is collected in the same environment we are trying to predict. However, even under ideal circumstances, the complexity of weather events can lead to prediction errors, especially over long prediction horizons.

By contrast, although human behavior may follow somewhat predictable patterns, human-based data, such as opinions and thoughts, is only observable indirectly by a person's actions or words and does not follow physical laws. Researchers sometimes try to augment their datasets using simulated data, but the effectiveness of this approach depends on the quality of the simulations employed.

Simulated data works only insofar as it is an accurate representation of reality. This is much easier to do when we understand the mechanics involved and can validate the simulations through testing, as in the case of physics-based systems such as missiles, tanks, planes, and ships. Simulation is far harder to do when we don't understand or cannot test the underlying mechanics, such as in social decision-making. In cases where we lack a clear comprehension of the mechanisms that connect inputs to outputs—as is true regarding almost any social or political question—simulated data may not be effective.

"Many intelligence reports in war are contradictory; even more are false, and most are uncertain."

- Carl von Clausewitz

Skewed data. Military commanders struggle to obtain accurate data about their own forces, let alone information about the enemy. For friendly forces, some units in the field can have poor communications and may not accurately transmit data about their position or status. Other platforms may flood the data stream with a single sensor in a particular area, causing humans and computers to over-rely on the single source. When it comes to adversaries, some are more adept at avoiding intelligence collection than others and some may be adept at active deception or data manipulation. Available data on some adversaries may be scarcer than for others.

Some information is more useful than no information, but biased data will skew an AI-DSS in a way that must be communicated to a user.[30] Furthermore, commanders should be especially wary of how biases in AI-enabled DSS might be amplified when they align with the personal or cultural biases of operators (for more, see next section on human-system integration).[31]

Human-based data presents different issues. Social media platforms often reflect a number of human biases: demographics may vary from one social media platform to another, and discourse on social media—which is shaped by technical, psychological, and social factors—may not reflect real-world behaviors or even opinions. Thus, using AI to predict violent uprisings on the basis of sentiment analysis of social media is an example of a problem where the underlying data is weak, and conclusions based in large part or entirely on this data should be treated with skepticism.

Some information is more useful than no information, but biased data will skew an AI-DSS in a way that must be communicated to a user.

Scarce data. While it is possible to gather large quantities of data on social media patterns, news reports, and market movements, the number of instances of the outcome to be predicted—uprisings, for example—can be extremely small (hundreds at most). Along the same lines, the military often faces a small-data problem because adversaries actively try to conceal valuable (and often rare) capabilities. In the absence of adequate observations, AI-based DSS are less valuable than traditional modes of intelligence analysis, which can combine insights and inferences that rely on a richer understanding of the relevant context in any particular case.

Human-System Integration

How does the human operator understand machine outputs? What are the capabilities and limitations of the human-machine team as a single system within a given context?

When evaluating whether it is appropriate to use an AI system in a decision support context, it is crucial to consider the properties of the AI system itself and how human operators are trained and prepared to interact with it. Three facets of human-system integration can pose important risks to the responsible use of DSS.

False expectations with LLMs. LLMs present distinct challenges from a human factors perspective. Notably, their propensity for confidently asserting factually incorrect information, often referred to as “hallucinations” or “confabulations,” can be highly misleading for users. Beyond hallucinations, LLMs exhibit further problematic behaviors, such as fabricating justifications for recommendations. These fabricated rationales tend to align with user expectations rather than reflect the true underlying reasoning process.[32] Research has unfortunately demonstrated that such unfaithful explanations can increase user acceptance of erroneous recommendations, and LLMs’ output must therefore be treated with caution for now.[33]

Human biases. Users have cognitive biases that can be exacerbated, as well as mitigated, through interactions with AI-DSS. These cognitive biases include, for example, confirmation bias (the tendency to seek supporting evidence for a hypothesis), ambiguity aversion (preferring known to unknown risks), and negativity bias (overweighting negative information over positive information).[34] One bias of particular concern with AI-DSS is automation bias, or assuming that an automated system is correct even when a user’s own senses indicate otherwise. The potential for humans to defer to algorithmic recommendations is well-documented and includes examples of tactical decision aids, like the Patriot and Aegis weapons systems.[35] Furthermore, stress in situations that involve the integration of large amounts of data can have a unique effect on the ability of humans to make decisions, as originally explored under the U.S. Navy’s Tactical Decision Making Under Stress project.[36] That AI-DSS might help humans make better decisions under stress or avoid harmful biases is a potential strength. To harness that potential takes awareness of the strengths and weaknesses of both AI and human operators in the design, deployment, use and maintenance of AI tools and accompanying organizational procedures.

“As a rule most men would rather believe bad news than good, and rather tend to exaggerate the bad news.”

- Carl von Clausewitz

Organizational biases. DSS can speed up decision-making and reduce the personnel required for certain tasks, but organizational assumptions about speed and human resources can lead to hasty or under-resourced decision-making and put pressure on military commanders to "do more with less." While some decisions may become faster as a consequence of AI and autonomous systems, the speed and quantity of decisions should be calibrated to contextual risks, as well as the quality of the decision-making processes and tools. It is important to determine where AI tools can augment operators and where they can truly reduce the number of necessary operators.

It is important to determine where AI tools can augment operators and where they can truly reduce the number of necessary operators.

Relatedly, the apparent ease of using some AI-DSS may predispose organizations and the people who work within them to use these systems in all cases, when in fact they may be best suited to only the most extreme situations when risk tolerance is high. One older example of aligning risk, technology, and organizational policy is the U.S. Navy's Close-in Weapons System. While CIWS is capable of firing autonomously, the autonomous setting is only enabled under strict conditions and following a well-known protocol. Similar risk-based policies and procedures should be established for the governance of AI-DSS.

Recommendations

AI-DSS could have strategic benefits in certain contexts, but those benefits are not without risks, which military commanders must work to control. With due consideration for current AI limitations and some of the guidelines already offered by, for example, DOD's Responsible AI Toolkit, we offer the following recommendations to leaders evaluating how to deploy these systems and mitigate their risks. We note that each of these recommendations has to do with organizations and people, not with the technologies themselves. Technical risk mitigation is important, and there are myriad efforts underway to technically improve AI systems. Here, however, we take as a premise that human governance actions will remain essential to mitigate current and future shortfalls of military AI-DSS and, moreover, that humans are the last line of defense for ensuring the employment of these systems is appropriate and done in compliance with international humanitarian law.

- **Set context- and risk-based criteria for deployment.** A commander deploying a DSS should take into consideration the time, place, and context of its application. Furthermore, deployment of a DSS should be reversible, and deployments should be adjusted or ended as factors on the ground change in ways that would affect AI performance. Commanders should establish a process and criteria for the continued deployment of an AI-DSS. That process must take into account the scope of the technology, the operational context, and also the readiness of the organization to properly leverage the DSS in accordance with rules of engagement with codified and tailored tactics, techniques, and procedures. In many ways, this approach is similar to how the military already handles commanders' guidance and special instructions, adjusting the authorization for the use of a system according to the tactical or strategic context.

- **Train and qualify AI-enabled DSS operators.** This includes any operator tasked with using an AI-based DSS, especially one supporting targeting or the employment of weapons. Training should start with study and experimentation in safe contexts but must build quickly from that base. A qualification regime for a DSS should have both schoolhouse and experiential components. Those DSS involved in lethal decisions could also be paired with rigorous, tailored examinations that lead to an official qualification designation for operators by a unit commander that is commensurate with their role in relation to the AI-DSS. The National Geospatial-Intelligence Agency's Responsible AI Training program is an early leader in designing and implementing a process for qualification and certification for users.[37]

- **Establish a continuous certification cycle.** Even with thoughtful design and user qualification, leaders should certify an organization's ability to faithfully execute a commander's intent using an AI-DSS. Because the environment in which a DSS is deployed is constantly changing, assessments should be frequent if not continuous. Since DSS are also repositories for data, logging and retaining information on the use and effectiveness of a DSS should be routine. Commanders should consider how to embed or share DSS performance metrics with trained data scientists, experts in AI test and evaluation, and operations analysts, both to validate the continued use of the system and to evaluate the effectiveness of a unit leveraging a DSS.

- **Designate a Responsible AI officer.** AI risks and opportunities are rapidly changing and show no signs of slowing. Moreover, there is a pressing need to establish broad-based AI literacy to help avoid some of the DSS risks we detail above. While there are unique factors to this challenge, there are similarities to other safety risks the military has addressed in the past. Depending on the mission, for example, military units often have occupational safety, weapons safety, or test safety officers. RAI officers could serve as a local, educated resource for operators as well as a conduit for reporting AI incidents and sharing new information from research or DOD-wide guidance organizations.

- **Document incidents and harms** from AI systems and maintain a central repository of reports to support knowledge sharing with developers, military analysts, and other users. AI systems have flaws, and human users will make mistakes. To best avoid repeating past mistakes, the Responsible AI Toolkit has made clear that AI harms should be documented.[38] As of publication, the DOD is still developing a documentation process specific to the military's needs, though several methods of reporting have been proposed for general use.[39] We support these efforts and would encourage the military to maintain a central repository of incident reports that can be analyzed and searched by operators and researchers to help avoid past mistakes and build trust through transparency. We would suggest that RAI officers take on the responsibility for documenting these harms in the same way that safety officers are often responsible for initial mishap reporting. Moreover, while military leaders may be predisposed to classifying or otherwise keeping all reports of incidents within DOD official channels, they should consider the value of sharing incident reports publicly. By making at least some reports public, the military may help prevent inadvertent harm in other militaries or in civilian contexts. Such transparency efforts could also build trust with the public. Making incident reporting a norm might also enable clearer communications between nations in the event an AI system fails in a way that might be misinterpreted as aggressive.

Conclusions

AI may improve the quality and speed of decision-making on the battlefield, but it cannot replace human judgment. The temptation to believe that AI-DSS are all-knowing is strong: They can draw on disparate data sources, integrate vast amounts of information, and generate recommendations at superhuman speeds. Moreover, their capabilities seem almost purpose-built for the unique challenges of war and the fog of battle. But AI-DSS also have critical weaknesses that require acknowledgment and intervention, and operators must beware of any automation bias they might harbor.

To seize the opportunity of AI-DSS in battle, commanders must prepare themselves and their teams to use them correctly. Questions of scope, data, and human-machine integration are important starting points for avoiding the weaknesses of these systems. These shortcomings alone are not necessarily reasons to avoid AI-DSS altogether, but these systems must be paired thoughtfully with human intelligence and judgment to achieve the best outcomes.

Appendix

As part of our research, we considered a number of AI-DSS that are being discussed, developed, advertised, requested, or used today. While the list isn't all-inclusive, it is intended to provide a representative snapshot of relevant systems, and the survey was used to inform this analysis.

- Proposal/Project Name: VISION; Organization: **Accenture**; Key Quote: "VISION (Versatile Intelligence System for Information and Operation Needs) empowers federal leaders to make informed, data-driven decisions on multifaceted issues such as conflicts, crises and disaster response by combining human judgment with models and simulations."[40]

- Proposal/Project Name: Lattice; Organization: **Anduril**; Key Quote: "Lattice streamlines the complexity of the decision-making process by presenting decision points—not noise—and using deep learning models to present recommended decision support to operators . . . Lattice enables real-time command and control over manned and unmanned assets across multiple domains, distributed geographies, and in contested communications environments."[41]

- Proposal/Project Name: multiple projects; Organization: **Clarifai**; Key Quote: "Rapidly turn mountains of data into plans of action for decision advantage to support the warfighter and the supply chain."[42]

- Proposal/Project Name: In the Moment; Organization: **DARPA**; Key Quote: "The Defense Sciences Office (DSO) at . . . DARPA is soliciting innovative research proposals for research and technology development that supports the building, evaluating, and fielding of algorithmic decision-makers that can assume human-off-the-loop decision-making responsibilities in difficult domains, such as combat medical triage."[43]

- Proposal/Project Name: DEFCON AI; Organization: **DEFCON AI**; Key Quote: "DEFCON AI, a next-generation software company building the foundation to enhance the modeling, simulation, and artificial intelligence enterprise across the Department of Defense (DoD), announced that it closed an [Air Force contract] to accelerate the transition from prototype to production code for DEFCON AI's operational-level logistics and mobility training software."[44]

- Proposal/Project Name: REASON; Organization: **IARPA**; Key Quote: "The Rapid Explanation, Analysis, and Sourcing Online (REASON) Program aims to develop technology that will enable intelligence analysts to substantially increase the quality of argumentation in their analytic reports through more effective use of evidence and reasoning."[45]

- Proposal/Project Name: The Gospel/Lavender/Where's Daddy;[46] Organization: **Israel Defense Forces**; Key Quote: "[Gospel] functions as a technical tool that fuses large amounts of data from across disparate datasets, and suggests to the intelligence analyst to focus his or her research on certain physical objects with greater potential of a military affiliation with the enemy . . . 'Lavender' is a general-purpose database that organizes and cross-references layers of several existing intelligence sources. It serves as a technical tool to help efficiently organize and connect data points relating to operatives in terror organizations in the Gaza Strip."[47]

- Proposal/Project Name: Wolf Howl; Organization: **Johns Hopkins University Applied Physics Laboratory**; Key Quote: "We'll give commanders the ability to 'wargame' different strategies from mission to unit within a given time frame or risk tolerance . . . That way, humans and computing machines can focus on aspects of planning they are each currently better suited to, and you really try to get the best of both worlds."[48]

- Proposal/Project Name: Global Planning and Monitoring (GLIMPS); Organization: **Leidos**; Key Quote: "The GLIMPS provides accurate, global forecasts on defined lead times of up to five years, focused on turbulent and complex environments, in order to provide the information needed to adapt to changing needs and resources. GLIMPS technology forecasts the effects of poverty, environmental degradation, political instability, and social tensions through big-data mining and machine learning of millions of open-source intelligence data points to discover the unseen relationships between indicators of stress and locations of potential instability."[49]

- Proposal/Project Name: ANTICIPE; Organization: **NATO Science & Technology Organization**; Key Quote: ANTICIPE "is designed to aid decision-making in an operational setting, using a built-in wargaming tool and advanced machine learning algorithms."[50]

- Proposal/Project Name: AIP; Organization: **Palantir**; Key Quote: "By leveraging the latent power of organizational data alongside interfaces for intelligent, fast decision making, AIP provides next-generation tooling."[51]

- Proposal/Project Name: Iris; Organization: **Rebellion Defense**; Key Quote: "Iris leverages state-of-the-art AI trajectory prediction methods to quickly find high-interest entities among cluttered environments for further investigation."[52]

- Proposal/Project Name: Donovan; Organization: **Scale AI**; Key Quote: "Synthesize insights in Donovan report templates. Quickly draft a course of action, briefing, or summary report. Capture mission critical information by prompting Donovan."[53]

- Project Name: Maven Smart System; Organization: **U.S. Army**; Key Quote: "The Scarlet Dragon version of MSS can access sensor data from diverse sources, apply computer vision algorithms to help soldiers identify and choose military targets, and then provide workflow support that enables a request to be approved by the chain of command in order to strike a target. It can also serve as a repository where battle damage assessments can be stored, as well as provide a map of the location of friendly forces and targets."[54]

Authors

Emelia Probasco is a senior fellow at CSET.

Helen Toner is the director of strategy and foundational research grants at CSET.

Matthew Burtell was a Horizon junior fellow at CSET, working on the foundational research grants team.

Tim G. J. Rudner is a nonresident AI/ML fellow with CSET and a faculty fellow at New York University.

Acknowledgments

For feedback and assistance, we would like to thank John Bansemer, Sam Bresnick, Lauren Lassiter, Larry Lewis, Arthur Holland Michel, Igor Mikolic-Torreira, Michael O'Conner, and Becca Wasser.

Document Identifier: doi: 10.51593/20240028

Endnotes

[1] Herodotus, *Histories*, trans. A. D. Godley (Cambridge, MA: Harvard University Press, 1920), book 1, chs. 54–77.

[2] DOD, *DOD Dictionary of Military and Associated Terms* (Washington, D.C.: DOD, November 2021), https://irp.fas.org/doddir/dod/dictionary.pdf.

[3] "ARCAS," Elbit Systems, accessed September 27, 2024, http://elbitsystems.com/product/arcas/.

[4] Seth Bonder, "Army Operations Research: Historical Perspectives and Lessons Learned," *Operations Research* 50, no. 1 (2002): 25–34.

[5] Richard Hillestad, Bart E. Bennett, and Louis R. Moore, "Modeling for Campaign Analysis: Lessons for the Next Generation of Models: Executive Summary" (RAND Corporation, January 1, 1996), www.rand.org/pubs/monograph_reports/MR710.html.

[6] Albert van Helden, "The Telescope in the Seventeenth Century," *Isis* 65, no. 1 (1974): 38–58, www.jstor.org/stable/228880; National Air and Space Museum, "The Blue Force Tracker System," Time and Navigation, Smithsonian Institution, accessed October 7, 2024, http://timeandnavigation.si.edu/multimedia-asset/the-blue-force-tracker-system; Paul K. Davis, "Capabilities for Joint Analysis in the Department of Defense" (RAND Corporation, 2016), www.rand.org/pubs/research_reports/RR1469.html.

[7] Gerardo Minguela-Castro, Ruben Heradio, and Carlos Cerrada, "Automated Support for Battle Decision Making: A Systematic Literature Review," *Military Operations Research Society* 27, no. 4 (2022): 5–24.

[8] For a case study on an AI-enabled DSS providing valuable inputs to military commanders; see Emelia Probasco, "Building the Tech Coalition: How Project Maven and the U.S. 18th Airborne Corps Operationalized Software and Artificial Intelligence for the Department of Defense" (CSET, August 2024), https://cset.georgetown.edu/publication/building-the-tech-coalition/.

[9] T. S. Allen, "Map Check – The Enduring Limitations of Russian C2" (working paper, 2024).

[10] Ryan Fedasiuk, Jennifer Merlot, and Ben Murphy, "Harnessed Lightning: How the Chinese Military Is Adopting Artificial Intelligence" (CSET, October 2021), https://cset.georgetown.edu/wp-content/uploads/CSET-Harnessed-Lightning.pdf.

[11] System Analysis and Studies Panel, "NATO STO Research Team Explores Ways to Support Decision Making with AI," NATO Science & Technology Organization, June 29, 2023, www.sto.nato.int/Pages/news.aspx.

[12] "ITM: In the Moment," DARPA, accessed July 12, 2024, www.darpa.mil/program/in-the-moment.

[13] Glenn Moy et al., *Recent Advances in Artificial Intelligence and Their Impact on Defence* (Canberra, Australia: Defence Science and Technology Group, April 2020), www.dst.defence.gov.au/sites/default/files/publications/documents/DST-Group-TR-3716_0.pdf.

[14] Graham Webster et al., "Full Translation: China's 'New Generation Artificial Intelligence Development Plan' (2017)," DigiChina, August 1, 2017, https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[15] Oleg Shakirov, "Russian Thinking on AI Integration and Interaction with Nuclear Command and Control, Force Structure, and Decision-Making" (European Leadership Network, November 2023), www.europeanleadershipnetwork.org/wp-content/uploads/2023/11/Russian-bibliography.pdf.

[16] 16 NATO, NATO Warfighting Capstone Concept (Norfolk, VA: NATO Allied Command Transformation, 2021), www.act.nato.int/wp-content/uploads/2023/06/NWCC-Glossy-18-MAY.pdf.

[17] 17 From DoD, Data, Analytics, and Artificial Intelligence Adoption Strategy (Washington, D.C.: DOD, June 27, 2023) document, released in November 2023, https://media.defense.gov/2023/Nov/02/2003333300/-1/-1/1/DOD_DATA_ANALYTICS_AI_ADOPTION_STRATEGY.PDF.

[18] NATO Science & Technology Organization, "Last year, the #NATO STO brought together scientists, researchers and military operators...," LinkedIn, March 25, 2024, www.linkedin.com/posts/natosto_nato-ai-activity-7178036772948340736-I4zD/; NATO Science & Technology Organization, "Using Artificial Intelligence to Enhance Military Decision-Making," YouTube, April 3, 2024, www.youtube.com/watch?v=A2ZAHrT3UwM.

[19] Fedasiuk, Melot, and Murphy, "Harnessed Lightning."

[20] In this analysis, decision support is different from physical control (like missile terminal guidance or swarm control for drones) to focus on those instances where a human is using an AI system to make a decision.

[21] Probasco, "Building the Tech Coalition"; "REASON: Rapid Explanation, Analysis and Sourcing Online," IARPA, accessed July 12, 2024, www.iarpa.gov/research-programs/reason.

[22] Naval Research Laboratory, "Recommender System for Enhancing Exploratory Data Analysis," TechLink, accessed February 25, 2025, https://techlinkcenter.org/technologies/real-time-3d-viewsheds-imaging-software/7088ca39-72cd-432e-b4db-3a35545db3b1.

[23] For example, see Deputy Secretary of Defense, *Implementing Responsible Artificial Intelligence in the Department of Defense*, May 26, https://media.defense.gov/2021/May/27/2002730593/-1/-1/0/IMPLEMENTING-RESPONSIBLEARTIFICIAL-INTELLIGENCE-IN-THE-DEPARTMENT-OF-DEFENSE.PDF; "RAI Toolkit: Overview," CDAO, accessed April 18, 2024, https://rai.tradewindai.com/;

"Responsible AI Guidelines," Defense Innovation Unit, accessed February 25, 2025, www.diu.mil/responsible-ai-guidelines.

[24] Arthur Holland Michel, *Decisions, Decisions, Decisions: Computation and Artificial Intelligence in Military Decision-Making* (Geneva: International Committee of the Red Cross, April 2024), https://shop.icrc.org/decisions-decisions-decisions-computation-and-artificial-intelligence-in-military-decision-making-pdf-en.html.

[25] Robert J. Doyle, Jr., "Real-Time Lines-of-Sight and Viewsheds Determination System," U.S. Patent 8,400,448, March 19, 2013; Palantir, "AI's Role in Reimagining the Classification System," *Palantir Blog*, February 22, 2204, https://blog.palantir.com/ais-role-in-reimagining-the-classification-system-2d7878c71217.

[26] Jacob Steinhardt and Helen Toner, "Why Robustness Is Key to Deploying AI" (Brookings Institution, June 8, 2020), www.brookings.edu/articles/why-robustness-is-key-to-deploying-ai/.

[27] Mazen El-Masri, Nabeela Altrabsheh, and Hanady Mansour, "Successes and Challenges of Arabic Sentiment Analysis Research: A Literature Review," *Social Network Analysis and Mining* 7, no. 54 (October 31, 2017), https://doi.org/10.1007/s13278-017-0474-x; Walaa Medhat, Ahmed Hassan, and Hoda Korashy, "Sentiment Analysis Algorithms and Applications: A Survey," *Ain Shams Engineering Journal* 5, no. 4 (December 1, 2014): 1093–1113, https://doi.org/10.1016/j.asej.2014.04.011.

[28] Tim G. J. Rudner et al., "Multi3Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery," *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 1 (2019), https://ojs.aaai.org/index.php/AAAI/article/view/3848.

[29] For short demonstration videos of the marketed systems, see Palantir, "Palantir AIP | Defense and Military," YouTube, April 25, 2023, www.youtube.com/watch?v=XEM5qz__HOU; Scale AI, "Scale Donovan | AI-Powered Decision-Making for Defense," YouTube, May 10, 2023, www.youtube.com/watch?v=0aBxvzdoyow.

[30] General Colin Powell made famous the 40-70 rule, implying that 40 percent was the minimum amount of information a military commander needed to make a decision about a situation and 70 percent was the maximum, as decisions made after 70 percent of the information was gathered were prone to be made too late to have an effect in time-sensitive scenarios.

[31] Michel, *Decisions, Decisions, Decisions*.

[32] Miles Turpin et al., "Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting," arXiv preprint arXiv:2305.04388 (2023), http://arxiv.org/abs/2305.04388.

[33] Gagan Bansal et al., "Does the Whole Exceed Its Parts? The Effect of AI Explanations on Complementary Team Performance," arXiv preprint arXiv:2006.14778 (2021), http://arxiv.org/abs/2006.14779.

[34] Tomáš Kliegr, Štěpán Bahník, and Johannes Fürnkranz, "A Review of Possible Effects of Cognitive Biases on Interpretation of Rule-Based Machine Learning Models," *Artificial Intelligence* 295 (June 1, 2021): 103458, https://doi.org/10.1016/j.artint.2021.103458.

[35] Lauren Kahn, Ronnie Kinoshita, and Emelia Probasco, "AI Safety and Automation Bias" (CSET, November 2024), https://cset.georgetown.edu/publication/ai-safety-and-automation-bias/.

[36] Rhona H. Flin, ed., *Decision Making under Stress: Emerging Themes and Applications*, repr. (Aldershot: Ashgate, 2005).

[37] Dr. Anna Rubinstein, interview with author.

[38] CDAO, "RAI Toolkit."

[39] Mia Hoffmann and Heather Frase, "Adding Structure to AI Harms" (CSET, July 2023), https://cset.georgetown.edu/publication/adding-structure-to-ai-harm/.

[40] "Generative AI for US Federal Agencies," Accenture, accessed February 25, 2025, www.accenture.com/us-en/services/us-federal-government/generative-ai.

[41] "Lattice OS," Anduril, October 2022, https://sldinfo.com/wp-content/uploads/2022/10/2022-slick-Lattice-OS-AUS.pdf.

[42] "AI for Military, Civilian, Intelligence, and State Agencies," Clarifai, accessed February 25, 2025, www.clarifai.com/government.

[43] DARPA, "In the Moment (ITM)," SAM.gov, accessed February 25, 2025, https://sam.gov/opp/baae2217401748dbaeb89a08044d6998/view.

[44] Loren Blinde, "DEFCON AI Wins Air Force SBIR Contract," *Intelligence Community News*, December 19, 2022, https://intelligencecommunitynews.com/defcon-ai-wins-air-force-sbir-contract/.

[45] IARPA, "REASON."

[46] Israel Defense Forces, "The IDF's Use of Data Technologies in Intelligence Processing," news release, June 18, 2024, www.idf.il/210062; Yuval Abraham, "'Lavender': The AI Machine Directing Israel's Bombing Spree in Gaza," *+972*, April 3, 2024, www.972mag.com/lavender-ai-israeli-army-gaza/; "Questions and Answers: Israeli Military's Use of Digital Tools in Gaza," Human Rights Watch, September 10, 2024, www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza; Elizabeth Dwoskin, "Israel Built an 'AI Factory' for War. It Unleashed It in Gaza," Washington Post, December 29, 2024, www.washingtonpost.com/technology/2024/12/29/ai-israel-war-gaza-idf/.

[47] Israel Defense Forces, "The IDF's Use of Data Technologies in Intelligence Processing."

[48] Ajai Raj, "Enabling Military Decision-Making at Operational Tempo," Johns Hopkins Applied Physics Laboratory, October 3, 2023, www.jhuapl.edu/news/news-releases/231003-enabling-military-decision-making-at-operational-tempo.

[49] "Products," Leidos, accessed May 9, 2024, www.leidos.com/products.

[50] NATO Science & Technology Organization, "Using Artificial Intelligence."

[51] "AIP for Defense," Palantir, accessed February 25, 2025, www.palantir.com/platforms/aip/defense/.

[52] "About Iris," Rebellion Defense, accessed February 25, 2025, https://rebelliondefense.com/products/iris.

[53] Scale AI, "Scale Donovan."

[54] Probasco, "Building the Tech Coalition."