8-5-2026

Fighting with Data: Design Implications for AI-Enabled Mission-Command Systems

C. Anthony Pfaff

SRAD
STRATEGIC RESEARCH & ASSESSMENT DEPARTMENT
US ARMY WAR COLLEGE

USAWC PRESS

MONOGRAPH SERIES

FIGHTING WITH DATA:
Design Implications for AI-Enabled Mission-Command Systems

DECISIVE POINT
The USAWC Press Podcast Companion Series

C. Anthony Pfaff

SRAD
STRATEGIC RESEARCH & ASSESSMENT DEPARTMENT
US ARMY WAR COLLEGE

The Strategic Research and Assessment Department (SRAD) answers strategic questions, informs strategic decisions, and expands expert knowledge supporting US Army and Joint Force requirements. The organization is divided into a Futures Research Group (FRG) that covers strategic competition, operational environmental security, emerging technologies, and AI integration and a Global Research Group (GRG) that includes the China Landpower Studies Center and researchers who study Europe, Eurasia, the Middle East, and the Western Hemisphere.

SRAD contributes to the war-fighting capability by identifying strategic challenges and opportunities, clarifying the associated risk, and reducing uncertainty regarding the best way to move forward. Through research, publications, conferences, and direct engagement with major commands and the think-tank community, SRAD translates military, global, and technological trends into actionable insights enabling effective responses. For operational units, this support improves readiness by informing doctrine, campaign planning, force design, and leader development. For strategic-level staffs, SRAD provides independent analysis that supports policy development, theater strategy, security cooperation, force posture decisions, and long-term modernization efforts.

By connecting scholarship, operational experience, and policymaking, SRAD serves as a strategic reconnaissance capability for the US Army—identifying emerging challenges, testing assumptions, and generating knowledge that improves the Army's ability to compete, deter, and, if necessary, win future conflicts.

SRAD is composed of civilian research professors, uniformed military officers, and a professional support staff, all with extensive credentials and experience. Researchers routinely engage organizations such as CENTCOM, the Joint Staff, Army Headquarters, combatant commands, and allied partners, providing expertise that helps decisionmakers align resources, priorities, and military capabilities with national objectives.

Fighting with Data: Design Implications for AI-Enabled Mission-Command Systems

C. Anthony Pfaff

August 2026

USAWC PRESS
US ARMY WAR COLLEGE

This publication is peer-reviewed. The views expressed in this publication are those of the author and do not necessarily reflect the official policy or position of the Department of the Army, the Department of War, or the US government. Authors of Strategic Research and Assessment Department and US Army War College Press publications enjoy full academic freedom, provided they do not disclose classified information, jeopardize operations security, or misrepresent official US policy. Such academic freedom empowers them to offer new and sometimes controversial perspectives in the interest of furthering debate on key issues. This publication is cleared for public release; distribution is unlimited.

******

******

Comments pertaining to this publication are invited and should be forwarded to: Director, Strategic Research and Assessment Department and US Army War College Press, US Army War College, 650 Wright Avenue, Carlisle, PA 17013-5049.

******

ISBN 1-58487-875-4

Cover Image Credits

Description: Image of a man made of blue energy.
Image by: USAWC Press on craiyon.com.
Date: June 18, 2026.
Website: https://www.craiyon.com/en.

Description: Vibrant and detailed view of JavaScript code on a screen.
Image by: Sabrina Gelbart on Pexels.
Date: Unknown, downloaded June 16, 2026.
Website: https://www.pexels.com/photo/full-frame-shot-of-abstract-pattern-249798/.

Foreword

Artificial intelligence is rapidly becoming central to how the US Army and the Joint Force sense, understand, decide, and act. Yet, the value of AI-enabled mission-command systems will depend on more than their ability to process greater volumes of data or accelerate staff procedures. As these systems assume a larger role in military decision-making processes by detecting patterns, generating explanations, and recommending courses of action, they will also change how military organizations produce knowledge and exercise judgment. The central challenge is therefore not simply technological. It is determining when commanders should trust machine-supported conclusions and how human judgment should be preserved within increasingly complex decision systems.

In this monograph, Dr. C. Anthony Pfaff offers a new framework for addressing that challenge. He argues that AI does not merely automate existing processes; rather, it redistributes cognition across a network of people, data, models, interfaces, and institutions. Drawing on military doctrine, cognitive science, and epistemology, he introduces an ecocognitive approach that evaluates AI-enabled systems according to how well the entire human-machine arrangement fits the operational environment. His proposed measures of representational, inferential, and adaptive fit provide military professionals with a practical way to assess whether these systems make the battlespace more intelligible, support sound reasoning under uncertainty, and improve decision quality through feedback and learning.

Fighting with Data is therefore both a study of trust and a guide to system design. Its findings have direct implications for commanders, staffs, acquisition professionals, doctrine developers, and those responsible for educating the future force. The monograph reminds us that decision advantage will not come from replacing military judgment with faster computation. It will come from designing human-machine systems that preserve responsibility, expose uncertainty, encourage critical inquiry, and adapt more effectively than the adversary. The Strategic Research and Assessment Department is pleased to offer this timely contribution to the US Army and Joint Force as they prepare for the demands of AI-enabled warfare.

US Army War College Press

Note: This text was created with AI assistance.

Executive Summary

Fighting with Data argues artificial intelligence (AI)–enabled mission-command systems do not simply make existing military decision-making processes faster; they fundamentally redistribute cognition across a human-machine network. In legacy systems, commanders and staffs rely on hierarchical, sequential processes in which humans manually collect, correlate, interpret, and act on information. In AI-enabled systems, machines assume a much larger share of the cognitive labor associated with sorting data, detecting anomalies, correlating signals, and generating candidate explanations. As a result, the role of the human shifts from assembling the picture to interpreting the picture's meaning, judging risks, and deciding how to act.

This shift in the human role creates challenges in two areas of AI-enabled mission command: trust and optimization. The first challenge concerns the limits of reliabilist approaches to machine output. In conventional settings, trust can be grounded in factors such as data curation, model validation and verification, robustness, calibration, and expert oversight. But these indicators are least helpful when AI output is future oriented—when systems are used to generate plans, predictions, or assessments of enemy intent whose truth cannot be known in advance and may remain contestable even after the fact. For example, a failed plan may still be the best one available under the circumstances, whereas a successful plan may rest on spurious correlations or opaque reasoning. Regarding optimization, most current approaches focus too narrowly on the relationship between human and machine, neglecting the larger distributed cognitive system through which machine output becomes knowledge and understanding. Thus, the real design challenge is not merely to optimize model performance. Instead, the challenge is to optimize the fit of the entire human-machine system to the operational environment.

This monograph frames this problem as one of epistemic design. The central question has shifted away from whether AI can produce outputs quickly and now asks whether the entire distributed system of humans, tools, data, and processes can produce knowledge and understanding that is trustworthy enough to support decision advantage in wartime. This issue is especially important because many AI systems, particularly generative and machine-learning models, operate probabilistically and often opaquely. These systems may offer plausible recommendations without making their underlying reasoning accessible to human users. Traditional approaches to trust based on reliability, past performance, or user experience are therefore insufficient on their own.

To explain the shift toward a focus on the trustworthiness of AI systems' outputs, this monograph contrasts legacy empiricist decision making with

AI-enabled reasoning. Legacy systems depend heavily on observation, induction, and deduction: Operators collect data, build confidence through corroboration, and apply rules to actions. Systems enabled by AI still use observation, induction, and deduction, but they are particularly valuable in supporting abductive reasoning—the generation and refinement of plausible explanations under conditions of uncertainty. In combat, where commanders must often act before ambiguity is resolved, the capacity to support abductive reasoning matters more than mere speed.

In this context, mission command should be understood as a cognitively distributed system. Knowledge does not reside solely in an individual commander, staff member, or analyst; rather, it emerges from interactions among people, interfaces, models, communications networks, and the operational environment. In practical terms, fighting with data means organizing mission command so data can move across echelons and functions to generate knowledge that allows the entire system to adapt to its environment more effectively. This organization can be facilitated by three connected data layers: foundational data, which provide strategic baselines; contextual data, which support operational understanding; and situational data, which connect tactical sensing and action to the broader system.

Because AI changes the way knowledge is produced, this monograph proposes an ecocognitive framework for trust. Rather than asking only whether a system reliably produces accurate output, this framework asks whether the system fits the operational environment well enough to help commanders generate useful understanding and adapt effectively. In this view, decision advantage comes from the entire system's ability to adapt faster than an enemy, not simply to decide faster. Whether a system sufficiently fits the operational environment depends on three broad functions: representation, inference, and adaptation. Accordingly, this study proposes metrics for representational fit, inferential fit, and output and adaptability fit to evaluate whether AI-enabled systems make the battlespace more intelligible, support the formation of plausible hypotheses, and improve decision quality over time through feedback.

Success in AI-enabled mission command will depend less on computational speed than on system designs that preserve human judgment, make uncertainty visible, sustain feedback loops, and maintain cognitive resonance between human users and machine outputs. In short, the future of war fighting lies not in replacing commanders with machines; rather, it lies in building human-machine cognitive ecologies that can learn, adapt, and fight effectively with data.

Fighting with Data: Design Implications for AI-Enabled Mission-Command Systems

Dr. C. Anthony Pfaff

Introduction

A sensor network detects an unusual dispersion pattern among a group of armored vehicles in a contested valley. At first glance, the formation looks routine, but a pattern-recognition model enabled by artificial intelligence (AI) flags it as anomalous. The vehicles are too widely dispersed for a defensive posture, yet they are also too irregularly dispersed for an offensive formation. Rather than dismissing the anomaly as noise, soldiers at the operations center hypothesize this new formation may indicate a feint or be protecting another high-value asset in the area, such as air defense, long-range fires, or command-and-control assets. To test these hypotheses, the soldiers adjust the model's parameters to search for additional contextual indicators such as electromagnetic signals, logistics-vehicle movements, or communication patterns consistent with either hypothesis. The model, without human prompting, assesses the anomaly against intelligence and operational reporting to determine whether any precedent exists. Meanwhile, the soldiers continue to collect sensor data from the area, and, as more evidence accumulates, the model updates its prior output, refining its analysis of the enemy's intent.

This example illustrates how AI-enabled mission-command systems will transform war fighting by redistributing cognition across human-machine networks, rather than by accelerating existing decision processes. This transformation will reframe military effectiveness as a problem of epistemic design—specifically, how well a distributed cognitive system produces knowledge commanders and staffs can use in decision-making through it's fit with the operational environment.[1] In the previous example, an initially unexplained anomaly—one that may have been dismissed prior to the advent of AI-enabled analysis—evolved into a process of hypothesis formation and refinement. This process was resolved by rapidly analyzing the broader environment for patterns that could

1. Mission command is "[t]he Army's approach to command and control that empowers subordinate decision making and decentralized execution appropriate to the situation." See Headquarters, Department of the Army (HQDA), *Mission Command: Command and Control of Army Forces*, Army Doctrine Publication 6-0 (HQDA, July 2019), x.

better explain the current observation, rather than by observing what the enemy formation did next. Non–AI-enabled targeting systems almost exclusively rely on empirical data and inductive logic, whereby observations—even when derived from sensors—yield general but fallible principles for understanding and interacting with the environment. The previously described process, on the other hand, captures the essence of abductive reasoning in an ecological setting in which goal-oriented actors compete: An unexpected observation gives rise to a plausible explanation, which is then iteratively tested and updated as actors converge on a more coherent understanding of the environment and the opportunities available to pursue their objectives.

The US Army and the Joint Force already have the foundation for effectively using these technologies. Current doctrinal publications and concepts of next-generation command and control already provide the context for understanding cognition and how it can be distributed to support effective decision making. But continued progress requires a clearer account of design—an account that gives humans reasons to trust the knowledge claims an AI-enabled system produces. Knowledge claims, in this context, are propositions about the environment that can be either true or false.[2] Cognition is the process by which agents acquire knowledge.[3] In classifier models, in which the system is used primarily to identify objects within a particular field, assessing trust can be straightforward in concept but difficult in practice. For example, if a classifier model identifies a tank, human operators can manually review the sensor data to confirm whether the model is correct.[4] Such systems lend themselves to a reliabilist approach, in which trust is a function of past performance.[5] In the case of large language models and generative AI (GenAI), an underlying reality against which to compare outputs often does not exist, or if it does, operators do not have access to it. For example, operators cannot know in advance whether an AI-developed course of action will work. Additionally, models that provide intelligence analysis of enemy actions typically operate with incomplete information, which makes assessing their output difficult. In such cases, trust rests almost entirely on how closely the output aligns with human experience.

As detailed later in this monograph, the problem for both approaches lies in the way metrics of reliability and experience create new conditions for failure

2. Linda Zagzebski, "What Is Knowledge?," in *The Blackwell Guide to Epistemology*, ed. John Greco and Ernest Sosa (Blackwell Publishing, 1999), 95.

3. Nagib Callaos, *Knowledge and Cognition* (International Institute of Informatics and Systemics, March 2013).

4. C. Anthony Pfaff et al., *Trusting AI: Integrating Artificial Intelligence into the Army's Professional Expert Knowledge* (US Army War College Press, February 2023), 31–33, https://press.armywarcollege.edu/monographs/959/.

5. Juan M. Durán, "Beyond Transparency: Computational Reliabilism as an Externalist Epistemology of Algorithms," preprint, arXiv, January 27, 2025, https://arxiv.org/html/2502.20402v1.

even as they address others. Reliabilist approaches can exacerbate shortcomings in human cognition, whereas experiential approaches can reinforce cognitive biases, thus increasing the likelihood of misinterpretations and missed opportunities. Thus, the central design question no longer concerns whether AI can produce answers quickly; rather, it concerns whether AI can help commanders and staff develop a more adaptable and usable understanding of the operational environment—one that enables decision dominance over an adversary. This emphasis on adaptability and usability suggests an ecocognitive approach that examines how cognitively distributed systems generate knowledge of an environment and provide advantages that enable goal achievement.

The use of the term "ecocognitive" here is deliberate, as it aptly illustrates the fact that fighting will entail both new concepts and new relationships among concepts to make full sense of the changes in war fighting AI adoption entails. Its use further illustrates to users the importance of understanding cognitive processes such as knowledge production for effective integration of machine processes and output into decision making. The effective application of AI will require users to think about cognitive processes differently than with non–AI-enabled systems. In such cognitive systems, commanders and their staffs still make judgments, take risks, and choose courses of action, with AI supporting these judgments, risks, and choices by shaping what people see, what they notice, and how they assess plausibility. The better the arrangement of people, tools, and data fits the operational environment, the more trustworthy the output.

The purpose of this discussion is to orient uniformed and civilian military practitioners—especially those who play a role in acquiring, integrating, and using AI-enabled technologies—to how the use of AI in decision-support systems, such as targeting and planning models, changes the way soldiers think about war fighting. Understanding these changes will enable acquisition efforts to interact more effectively with AI engineers, scientists, and commercial providers in optimizing system design. Drawing on insights from ecocognitive epistemology, which views knowledge as emerging from dynamic interactions among cognitive agents, their tools, and their environment, this monograph develops a framework for evaluating trust in AI-enabled systems that complements reliabilist and experiential approaches. The redistribution of cognitive tasks between humans and machines alters knowledge production and judgment, necessitating a metric of fit by which one can assess whether AI-enabled systems enhance or undermine military decision making. Whereas this discussion focuses on US Army mission-command and targeting systems, many of its findings and recommendations also apply to other decision-support systems.

The discussion is organized into five main sections. The first section, "Fighting with Data," contrasts legacy, empiricist approaches to battle command and targeting with AI-enabled processes that redistribute cognitive labor between humans and machines, reframing military decision making as a problem of epistemic design rather than better automation. The second section, "The Nature of Knowledge," establishes the epistemological foundations for understanding how knowledge is produced and validated in cognitively distributed systems, contrasting legacy approaches with the framework proposed in this monograph. The third section, "Cognitive Distribution, Mission Command, and the Targeting Process," examines how distributed cognition manifests in the targeting ecosystem and maps how knowledge and decision authority are shared among human and nonhuman agents.

Drawing on the insights of the first three sections, the fourth section, "Epistemology and Design," develops a revised ecocognitive model for evaluating AI-enabled systems, identifying fit and cognitive resonance as design standards for epistemic trust. In this model, approaches to design epistemology serve two functions: determining whether a design is suitable for a particular environment and establishing the basis on which output may be trusted. The fifth section, "Designing Measures of Fit," translates these conceptual insights into metrics that are representational, inferential, and adaptive, guiding the design, assessment, and governance of AI-enabled targeting systems. Together, these sections provide both a theoretical framework and a set of design principles, ensuring AI technologies enhance, rather than simply displace, the human capacity for understanding in war.

Fighting with Data

The most effective AI-enabled systems do not simply accelerate the military decision-making process; they redistribute it. For example, according to a 2024 report by Georgetown University's Center for Security and Emerging Technology, the integration of AI within the XVIII Airborne Corps enabled the organization's fire support element to accomplish with 20 people what had previously required significantly more people during the Iraq War.[6] Such a dramatic reduction in human requirements was not simply a function of AI's ability to accelerate decision making; rather, it stemmed from how AI redistributes the cognitive load throughout the mission-command process. Given how profound this redistribution was, members of the XVIII Airborne Corps' fire support element often described

6. Emelia S. 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* (Center for Security and Emerging Technology, August 2024), 4.

Fighting with Data

the experience as "fighting with data," rather than with the weapons and other assets that once executed war fighting.[7]

This experience—which the author of this monograph observed evolving over four years of the XVIII Airborne Corps' experiments with AI-enabled targeting systems—suggests fighting with data entails more than data literacy. It also requires a data-centric, adaptive, and collaborative approach that will transform how military organizations convert raw data into informed judgment.[8] Understanding the significant differences between weapon-focused, legacy processes and data-driven processes begins with understanding how legacy and AI systems navigate the data–information–knowledge–understanding model described in Army Doctrine Publication 3-13, *Information*, which explains how data are transformed into human judgment. In this model, data consist of unprocessed observations that become information when put into context. Information becomes knowledge when analyzed for operational implications; and knowledge becomes understanding when applied to comprehend a situation's inner relations in a way that enables decision making and drives action.[9]

To understand how AI-enabled systems can transform war fighting, one must also understand what these systems are replacing and how they are doing so. Because the purpose of these systems is to produce knowledge of the operational environment that enables commanders to act, understanding how to assess knowledge claims becomes necessary, placing the problem within epistemology—the study of what knowledge is, how it can be justified, and when belief is warranted.[10] Legacy military decision-making systems, defined here as non-AI-enabled systems or processes, typically exhibit deliberate, top-down processes that rely more on individual experience than on data in determining whether judgments are sound.[11] In such systems, knowledge is grounded in an empiricist epistemology, in which human agents rely on direct perception and form judgments about those perceptions through inductive and, sometimes, deductive reasoning—a topic discussed later.

Under a legacy system, the process described in the opening example would look very different. Without AI, the processes of anomaly detection and hypothesis generation would be slower and more fragmented, because humans

7. Joseph O'Callaghan, "Fighting with Data," unpublished monograph, 2024.

8. O'Callaghan, "Fighting with Data," 6–8.

9. HQDA, *Information*, Army Doctrine Publication 3-13 (HQDA, November 2023), 1-2; and O'Callaghan, "Fighting with the Data," 6.

10. Robert Audi, *Epistemology: A Contemporary Introduction to the Theory of Knowledge* (Routledge, 1998), 1–5.

11. Jason N. Adler, "Modernizing Military Decision-Making: Integrating AI into Army Planning," *Military Review* (August 2025); and Richard Hart Sinnreich, "Variables and Constants: How the Battle Command of Tomorrow Will Differ (or Not) from Today's," in *Battle of Cognition: The Future Information-Rich Warfare and the Mind of the Commander*, ed. Alexander Kott (Praeger Security International, 2008).

would have to review sensor data manually, correlate disparate intelligence sources, and recall relevant patterns from experience rather than through rapid, parallel pattern recognition. As a result, the targeting cycle would unfold more slowly and sequentially, increasing the likelihood of soldiers missing or misinterpreting anomalies. Other indicators, such as electromagnetic signals, patterns of logistics movements, or radio communications, would have to be addressed manually by separate, often stove-piped, staff sections. Correlating these data would require time-consuming cross talk, manual deconfliction, and iterative briefings, giving an enemy time to emplace whatever assets it was moving.[12]

The legacy system described in this discussion is an intentional straw man, serving as a baseline for understanding AI-enabled change. Whereas many legacy features persist in current targeting and command-and-control systems, the Army and the Joint Force have made progress in integrating new technologies and enhancing system capabilities. Moreover, the Chief of Staff of the Army and the Army Futures Command (now part of the US Army Transformation and Training Command) have articulated a comprehensive vision for integrating and applying new technologies to enable decision dominance. In a white paper titled *Army Multi-Domain Transformation: Ready to Win in Competition and Conflict*, the Chief of Staff of the Army defined decision dominance as the ability to sense, understand, decide, and act faster than an adversary. Such dominance is enabled by a "convergence" of technologies that provides "the ability to see, sense, communicate, shoot, and move at speed and scale" by connecting sensors to shooters through the right command-and-control nodes.[13]

Army Futures Command Pamphlet 71-20-9, *Army Futures Command Concept for Command and Control 2028: Pursuing Decision Dominance*, provides a path toward realizing the chief of staff's vision. Importantly for the present discussion, this solution is described as a distributed cognitive network composed of people, communication networks, processes, and a command-post constellation, enabling both projection into and adaptation to the operational environment. This network is intended to enable commanders to maintain decision advantage by engaging in a cycle of sensing and understanding, deciding, acting, and assessing.[14] Moreover, the network learns and adapts to individuals and

12. Joshua Thayer, "The AI Revolution in Military Leadership: Augmenting, Not Replacing, Human Judgment," *Medium* (blog), December 1, 2024, https://medium.com/@grizzlythayer/the-ai-revolution-in-military-leadership-augmenting-not-replacing-human-judgment-5c808200b12c.

13. HQDA, *Army Multi-Domain Transformation: Ready to Win in Competition and Conflict*, Chief of Staff Paper no. 1 (HQDA, March 2021), 8–9. The author owes this point to an anonymous reviewer.

14. Army Futures Command (AFC), *Army Futures Command Concept for Command and Control 2028: Pursuing Decision Dominance*, AFC Pamphlet 71-20-9 (AFC, July 2021), 18. The author owes this point to an anonymous reviewer.

Fighting with Data

organizations to optimize cognitive processes.[15] The pamphlet also describes features—including effective human teams, consistent processes, resilient networks, reliable data, and explainable outputs—that enable the network to fit better into dynamic operational environments.[16] Also identified are engineering solutions, such as semantically adaptive network control, which modifies applications and network resources according to specific operational environments and feedback.[17] The Army and the Joint Force's efforts show military professionals will have to understand how AI-enabled systems generate and evaluate knowledge.

To operationalize the systems described in US military documents, one must first understand how the systems' application affects the war-fighting experience. As observed in corps-level targeting processes, fighting with data produces a war-fighting experience very different from that of the legacy approach, treating information as a form of combat power that can be entangled, massed, and applied across all levels of war to provide decision advantage. Rather than relying on hierarchical, sequential decision making, "fighting with data" distributes the cognitive load of targeting across a network that connects what can be referred to as "situational," "contextual," and "foundational" data—layers roughly corresponding to the tactical, operational, and strategic levels of war.[18] This approach allows for the rapid transformation of raw data into understanding, reframing war fighting as an adaptive, data-centric enterprise in which cognitive capacity is expanded by technology and in which decision making is a continuous, collaborative process.[19]

In an AI-enabled process, machines handle a greater share of the sorting and correlation required before information reaches the staff. As illustrated by this monograph's opening example, these machines can classify imagery, detect anomalies, identify likely matches, surface patterns across time and space, and, depending on the system, generate candidate explanations or courses of action. This machine involvement does not remove humans from mission command; it changes the human role. The staff spends less time assembling the picture and more time testing the picture's meaning, deciding what matters, and determining how much risk to accept.

This shift is easiest to see through the well-known observe–orient–decide–act loop. Developed by US Air Force Colonel John Boyd, the loop enables effective

15. AFC, *Command Concept*, 28.
16. AFC, *Command Concept*, 29.
17. AFC, *Command Concept*, 57.
18. O'Callaghan, "Fighting with Data," 7–9.
19. O'Callaghan, "Fighting with Data," 7–9.

reactions to a rapidly changing environment.[20] Originally intended as a tool to reform the design of fighter aircraft, the loop prescribes a sequence that begins with collecting data, which are analyzed to orient actors to action-relevant features of the environment. Because action changes the environment, the loop repeats itself endlessly, or at least as long as actors need to gain or maintain the initiative.[21] Artificial intelligence (AI) can impact observation by expanding how much data a formation can detect and sort, and it can change orientation by altering how quickly the staff can connect observations to context. The second point is at the heart of the issue: Orientation is the step during which commanders and staffs make sense of the fight, connecting data to operational purpose, adversary behavior, terrain, doctrine, and timing. Whereas AI can support this work, it can also distort it if the model is brittle, the data are biased, or the staff treats probability as certainty.[22]

Boyd also saw the design problem for military systems in epistemic terms, though he did not use the term. In his essay, "Destruction and Creation," he analogized that military organizations were much like living things: goal-oriented organisms that desire to survive and grow under conditions of scarce resources. In such conditions, especially where organisms compete, survival and growth depended as much, if not more, on adaptability than strength or speed. For Boyd, that adaptability depends on the constant destruction and creation of mental models that allow the organism to make sense of its environment. In his view, while the new mental model may better match reality in the moment, eventually it will drift as internal inconsistencies and entropy create a gap between model and reality.[23]

John Kay and Mervyn King's distinction between puzzles and mysteries can aid an understanding of how humans and machines can distribute cognitive tasks to better close that gap. Puzzles have a right answer that can often be improved with more data and processing power.[24] Battlefield problems such as identifying a vehicle type, sorting sensor tracks, matching signals to known emitters, or detecting a pattern in fire support are examples of puzzles. As described earlier, AI is already an exceptionally valuable part of executing such tasks. Other battlefield problems

20. Christopher Tuck, *Understanding Land Warfare* (Routledge, 2014), 99.

21. Lawrence Freedman, *Strategy: A History* (Oxford University Press, 2013), 196–97.

22. Emelia S. Probasco et al., *AI for Military Decision-Making: Harnessing the Advantages and Avoiding the Risks* (Center for Security and Emerging Technology, April 2025), 1–2, 20–22; and US Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (Department of Defense, February 2019).

23. John Boyd, "Destruction and Creation," A Discourse on Winning and Losing, Colonel John Boyd Archive, September 3, 1976, https://www.coljohnboyd.com/#pdf-destruction-and-creation.

24. John Kay and Mervyn King, *Radical Uncertainty: Decision-Making Beyond the Numbers* (W. W. Norton and Company, 2021), 20.

Fighting with Data

are mysteries, involving intent, deception, shifting priorities, enemy adaptation, and the operational significance of the fight. Access to larger volumes of data does not automatically enable commanders to solve mysteries, and commanders remain responsible for making judgments about mysteries.[25] Whereas AI can support commanders by surfacing possibilities, contradictions, and other relevant patterns, it does not remove the radical uncertainty mysteries entail. Mysteries may have a solution, but telling when one has found that solution can be difficult, if not impossible.

The distinction between puzzles and mysteries clarifies the division of labor required by mission command. Whereas humans understand, judge, prioritize, and accept responsibility, AI systems compute, predict, classify, and perform pattern matching. The whole system performs cognitive functions because people, tools, and processes work together to create understanding. Fighting with data, therefore, means organizing in a way that allows data to move across domains, echelons, and functions without losing meaning. In practice, this transition requires at least three connected layers of information: foundational, contextual, and situational (see figure 1).[26]

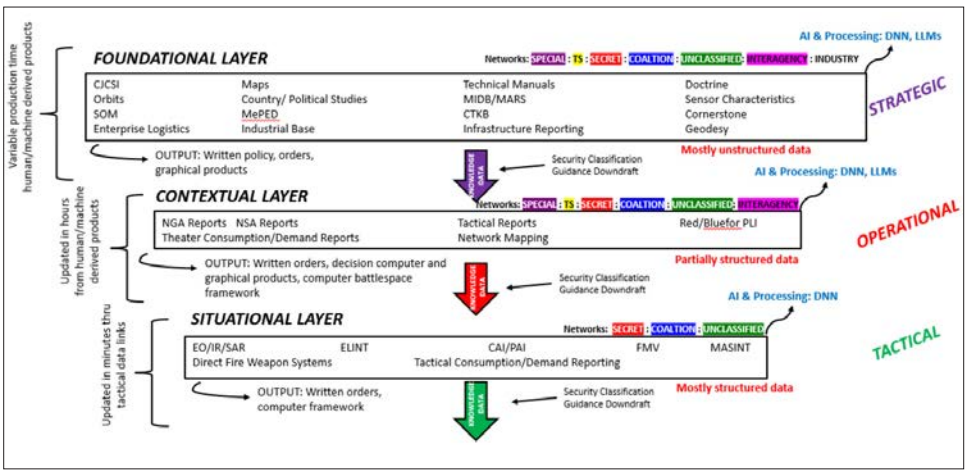

Figure 1. Data layers
(Source: Joseph O'Callaghan, "Fighting with Data," unpublished monograph, 2024.)

Foundational data provide the strategic baseline, defining the physical world, the resourcing, and the policy frameworks that anchor military and interagency understanding. These datasets are managed within large-scale data-management platforms, constituting a layer that includes national-level geospatial information, strategic intelligence, logistics networks, contracts, and policy documents.

25. Kay and King, *Radical Uncertainty*, 21.
26. O'Callaghan, "Fighting with Data."

Once the data are ingested into a data-management platform, ontologies—defined as the set of concepts and categories that establish relationships between information and data—can be used for data entanglement.[27] Such entanglement involves connecting the data in a way that allows them to be modified as new information arrives.[28] In this context, foundational data establish the basis on which all other knowledge generated by the system rests. The data at this level change over longer periods than at other levels, serving as a common reference point for decision making.[29]

Supported by a common ontology and entangled data, contextual data sit at the operational level, where commanders gain an understanding of the operational environment. This level is also where commanders and staffs translate strategic intent into operational design by projecting data onto the foundational level to support resourcing, sensing, shaping, and synchronization. Enabling an understanding of the operational environment requires integrating operational, intelligence, sustainment, and publicly available information to speed up workflows at scale.[30] As discussed in later sections, the resulting understanding at the operational level allows commanders to assess how well mission-command systems fit the fight. Because of its importance to environmental fit, this data layer serves as the center of gravity for a cognitively distributed battle-command system.

The situational layer links contextual understanding to mission requirements. The decide–detect–deliver–assess (D3A) cycle exemplifies how contextual understanding is integrated with sensor data and other tactical-level information to inform commanders' decisions. Situational data connect directly to sensors, weapons, and operators, ensuring decisions are informed by, and feed back into, the larger system. At this level, data serve as the feeder for selected weapon systems. The challenge at this level is to process large volumes of situational data to enable human understanding of them, given what is known at the foundational and contextual layers.[31]

27. Charu C. Aggarwal, *Neural Networks and Deep Learning: A Textbook* (Springer Nature Switzerland, 2023), 24.

28. Mark Rasch, "Data Entanglement, AI and Privacy: Why the Law Isn't Ready," *Security Boulevard* (blog), February 25, 2025, https://securityboulevard.com/2025/02/data-entanglement-ai-and-privacy-why-the-law-isnt-ready/; and James Aspnes et al., "Towards a Theory of Data Entanglement," in *Computer Security – ESORICS 2004*, ed. Pierangela Samarati et al. (Springer, 2004). Entanglement was originally developed to promote resistance to tampering by ensuring the modification of some data would require modifying related data, thereby making tampering more easily detectable. But as O'Callaghan ("Fighting with the Data," 7–8) notes, entanglement can enable faster, more productive access to data, improving output.

29. O'Callaghan, "Fighting with Data," 9–14.

30. O'Callaghan, "Fighting with Data," 20–28.

31. O'Callaghan, "Fighting with Data," 28–33.

Fighting with Data

These layers are integrated through object-based production and managed via data meshes and fabrics, producing a self-reinforcing system in which data flow both vertically and horizontally. Object-based production enables that flow by creating a conceptual object (such as a person, a place, or a thing) that acts as a "bucket," allowing information from every layer to converge and paint the most complete picture possible, given what information is known.[32] Foundational data ground situational awareness in verified geospatial and policy realities; contextual data provide the scaffolding that links events and objectives; and situational data inject immediacy and feedback, enabling adaptation and learning. Together, these layers transform the mission-command process from a linear chain of reports and decisions into a distributed cognitive ecosystem in which each level contributes to and draws from a common, evolving knowledge base.

Entangling data so they can be accessed within and between layers allows commanders to, as the Army's Office of the Chief Information Officer describes, apply the "right data, at the right time, at the right place," to "out-think and out-pace any adversary."[33] This description is one way to think about what massing data means. In the process described previously, massing data requires collecting them from a variety of sources and sensors, fusing them into a coherent picture, curating them for reliability, analyzing them using appropriate models, and distributing them quickly to the right organization at the right echelon to enable action. Data, however, is not massed by accumulation alone. Rather it must be structured and aligned to detect the opportunities the interaction between friendly and enemy forces in the operational environment affords.[34] This massing is essentially the logic behind the Combined Joint All-Domain Command and Control, which is not a single platform, but better thought of as a set of interconnected capabilities supporting informed decision making.[35]

The military value of massing data comes from turning dispersed observations into actionable knowledge faster than the adversary can adapt. For example, an AI-enabled targeting system might mass foundational data on friendly and enemy doctrine, contextual data on friendly and enemy operational capabilities and

32. O'Callaghan, "Fighting with the Data," 16; and Catherine Johnston et al., "Transforming Defense Analysis," *Joint Force Quarterly* 79 (4th Quarter 2015): 12–18.

33. Office of the Chief Information Officer, *Army Data Plan* (US Army, 2022), 2.

34. S. Tucker Browne, "Rethinking Technological Readiness in the Era of AI Uncertainty," preprint, arXiv, April 15, 2025, https://arxiv.org/html/2506.11001v1; and Benjamin Jensen and Jake S. Kwon, "The Convergence Conundrum: Achieving Mass in the Era of Artificial Intelligence," Modern War Institute at West Point, April 1, 2025, https://mwi.westpoint.edu/the-convergence-conundrum-achieving-mass-in-the-era-of-artificial-intelligence/.

35. "CJADC2," Chief Digital and Artificial Intelligence Office, n.d., accessed May 27, 2026, https://www.ai.mil/Initiatives/CJADC2/; and US Government Accountability Office, *Defense Command and Control: Further Progress Hinges on Establishing a Comprehensive Framework*, GAO-25-106454 (Government Accountability Office, April 2025).

constraints, and situational data from all-source intelligence feeds, current weather and terrain, and command priorities to generate hypotheses, prioritize collection, recommend options, and update the assessment after effects are observed.[36] In that sense, the force is not just massing data; it is massing cognitive capacity across humans, sensors, models, networks, and decision processes.

To varying degrees, the Army is already using these concepts and processes in mission-command systems. The Next Generation Command and Control program, for instance, integrates AI-enabled systems to improve cross-layer integration and build increasingly complex, distributed cognitive networks.[37] For example, through the Ivy Sting series of exercises, the program has increased the number of networked firing systems and command-and-control nodes from one system paired with one node to three systems paired with six nodes. Future exercise iterations will involve adding sensors and various other capabilities, such as logistics and casualty-evacuation capabilities.[38] Whereas these innovations aim to reduce the cognitive burden on humans, they do so by redistributing cognitive functions they previously performed. This development introduces complexity that must be addressed to establish the conditions for trust. Addressing this complexity requires an understanding of the epistemic challenges AI integration poses for cognitively distributed systems.

Epistemic Challenges in Cognitively Distributed Systems

In cognitively distributed systems, knowledge is not simply the aggregation of individual cognitive agents; rather, it emerges from interactions among agents, tools, and the environment. Knowledge is thus externalized, encoded in the system and the environment in which it operates.[39] Consider, for example, a pilot's checklist. Certified pilots certainly know how to fly a plane, but they also are cognitively incapable of doing so without some form of cognitive support, such as a checklist, that helps them bear the cognitive load.[40] Whereas this externalization of knowledge may be present in any system composed of cognitive agents—AI enabled or not—

36. Wyatt Hoffman and Heeu Millie Kim, *Reducing the Risks of Artificial Intelligence for Military Decision Advantage* (Center for Security and Emerging Technology, March 2023), 5–7, 13.

37. Mark Pomerleau, "How the Army Built Next-Gen Command and Control," *DefenseScoop*, March 20, 2025, https://defensescoop.com/2025/03/20/how-army-built-next-gen-command-and-control-ngc2/.

38. "Next Generation Command and Control: Ivy Sting 2 Media Roundtable," U.S. Army, November 5, 2025, https://www.army.mil/article/289180/next_generation_command_and_control_ivy_sting_2_media_roundtable.

39. Andy Clark and David Chalmers, "The Extended Mind," *Analysis* 58, no. 1 (January 1998): 7–10.

40. Michael Maddox, *Human Factors Guide for Aviation Maintenance and Inspection* (Federal Aviation Administration, n.d.), 8; and James Hollan et al., "Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research," *ACM Transactions on Computer-Human Interactions* 7, no. 2 (June 2000): 176.

it implies AI-enabled systems require a different metric for assessing knowledge and judgment claims.

Further contributing to the opacity of AI-enabled systems is their potential to introduce another cognitive agent—beyond humans—that mediates human perception and redistributes cognitive load in ways that are neither transparent nor always understandable. This redistribution impacts how tasks and associated knowledge are managed across different components of a process. In redistributing the cognitive load, AI also introduces new challenges related to bias, control, and agency redistribution, affecting systems and processes that rely on legacy epistemologies and logic to account for trust. This lack of transparency is exacerbated in combat operations—in which pressures are high and humans are often already cognitively impaired—diminishing humans' capacity to vet output from AI-enabled systems.[41] Such conditions can lead to undesirable outcomes and undermine trust in judgments supported by AI-enabled systems.

The opacity of AI technologies also complicates the assessment of the cognitive functions an AI agent performs and the portion of the cognitive load the agent shares.[42] Unlike conventional tools, an AI agent is not a passive participant in the cognitive process. Instead, the agent's ability to simulate learning, assess output, and adjust its programming allows it to act as an "eco-cognitive engineer," contributing to the creation of cognitive niches that optimize system performance.[43] Whereas AI agents do not replicate or even mimic human cognitive processes, they do displace them. In this displacement, the agents relocate functions associated with knowledge production—such as reasoning, interpretation, and judgment—from human operators to processes that are only partially transparent or controllable. This relocation complicates the operators' ability to assess knowledge claims, affecting their capacity to establish accountability and trust in AI-enabled systems.

Opacity is not the only, or even the most important, epistemic challenge. The output of AI, especially GenAI models, is probabilistic. For example, GenAI models such as ChatGPT calculate the probability that a word token will follow, given what preceded it, starting with the user's prompt.[44] In such systems, models getting so much right in their responses is more remarkable than their getting anything wrong. Thus, the temptation to use GenAI to shorten military

41. Tuck, *Understanding Land Warfare*, 24–26.

42. Ben Chester Cheong, "Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making," *Frontiers in Human Dynamics* 6 (July 2024): 1–11.

43. Tommaso Bertolotti, *Patterns of Rationality: Recurring Inferences in Science, Social Cognition and Religious Thinking* (Springer, 2015), 172.

44. Stephen Wolfram, "What Is ChatGPT Doing . . . and Why Does It Work?," *Stephen Wolfram Writings* (blog), February 14, 2023, https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/.

decision cycles is understandable. Humans frequently produce outputs, such as staff estimates needed for planning, that are incorrect and require validation. Therefore, skipping the slower, human-driven development process—and retaining human validation when suitable GenAI models are available—makes sense.

But this approach may be appropriate only in cases in which GenAI models meaningfully replicate the human computational and geospatial capabilities required for accurate planning. Often, this condition does not hold. For example, mission-analysis programs, such as the planning-support AI platform Scale Donovan, offer to "[p]ut advanced AI capabilities in the hands of warfighters when and where they need it most."[45] But Scale Donovan, like most such systems, lacks computational and geospatial capabilities. As a result, system output depends on the model's sophistication and the comprehensiveness of the available data. Thus, if asked to account for terrain or movement rates, the model would guess an answer rather than perform the calculations itself.[46]

Nevertheless, given a sufficiently sophisticated model and robust data, the basis for claiming a probabilistic response is inferior to a computational one remains unclear. This challenge is well illustrated by Orson Scott Card's *Ender's Game*, in which Andrew "Ender" Wiggin, a boy genius, can discern more patterns, account for more factors, and develop better (though frequently counterintuitive) solutions than others using traditional approaches.[47] Ender's capability is precisely what AI-enabled systems are designed to do. In fact, as Andrew Hill and Dustin Blair point out, human history is filled with examples of counterintuitive judgments being superior to those based on generally accepted principles and calculations. They relate the story of the Vicksburg Campaign, in which Union General Ulysses S. Grant rejected the generally accepted principles that emphasized secure supply lines and safe bases of operations, instead relying on the concentration of force to outnumber and defeat the enemy. Grant left his supply lines behind, maneuvered between two enemy forces, and crossed the Mississippi River with an unprotected line of retreat.[48] Despite carrying risks, Grant's move succeeded and Vicksburg fell. Grant's decision may be better understood as a gamble that was motivated by political pressure and that aimed to avoid defeat, rather than as a product of brilliant abductive reasoning. But Grant

45. "Donovan," Scale AI, n.d., https://scale.com/donovan.

46. A. Blair Wilcox and C. Anthony Pfaff, "Responsibly Pursuing Generative Artificial Intelligence (GenAI) for the War Fighter," *Parameters* 55, no. 4 (Winter 2025–26): 7–13, https://press.armywarcollege.edu/parameters/vol55/iss4/3/; and William J. Barry and Aaron "Blair" Wilcox, *Centaur in Training: US Army North War Game and Scale AI Integration*, Issue Paper 2–25 (Center for Strategic Leadership, US Army War College, August 2025).

47. Orson Scott Card, *Ender's Game* (Tom Doherty Associates, 1994).

48. Andrew Hill and Dustin Blair, "Alien Oracles: Military Decision-Making with Unexplainable AI," War on the Rocks, September 26, 2025, https://warontherocks.com/2025/09/alien-oracles-military-decision-making-with-unexplainable-ai/.

nevertheless saw possibilities others—including his subordinate, General William Tecumseh Sherman—did not.[49] Moreover, Grant and Sherman had no independent metric or set of experiences by which to assess the decision's chance of success. If, as Hill and Blair contend, AI will increasingly place commanders and staffs in similar situations, developing such a metric seems imperative.

Overcoming such epistemic challenges requires a different framework for understanding teaming between AI and humans—one that accounts for how the entire cognitive system, including humans, tools, and models, interacts with the environment in which the system is designed to operate. Here, the ecocognitive approach to resolving epistemic concerns can be particularly helpful. This approach is grounded in nature, whereby an organism's survival depends on the fit between its cognitive capacities and its environment. Applied to combat operations, this naturalized approach opens up different ways of thinking about knowledge and trust. By treating combat organizations as goal-oriented organisms that seek to survive and thrive in a particular environmental context, knowledge claims can be assessed using a metric of fit that aligns cognitive processes with ecological demands. Fit is achieved when an organism's or organization's cognitive capacities clearly represent relevant data and enable hypothesis generation, testing, and refinement. These processes must support effective, adaptive responses to environmental factors—such as predators and adversaries—that threaten an organism's or organization's ability to survive.

The Nature of Knowledge

Before answering the question of how one can trust knowledge produced in a cognitively distributed system, one must first be clear about what standards factual or other propositional claims must meet to be considered knowledge. Epistemologists typically distinguish among three kinds of knowing: knowing-that, knowing a person or thing, and knowing-how. Knowing-that refers to knowledge of facts—for example, a targeteer knowing a particular image is that of a tank. Knowing a person or thing implies understanding, awareness, or familiarity, which may involve claims of knowing-that and knowing-how aggregated for greater meaning. For example, claims such as "he knows the commander" may mean a person recognizes the commander, knows facts about the commander's personality, or knows what the commander would do in a particular situation. Knowing-how refers to one's ability to do something, such as speak a foreign language or, more salient for the present discussion, detect, prioritize, and engage targets on the

49. Allan R. Millett et al., *For the Common Defense: A Military History of the United States from 1607 to 2012* (Free Press, 2012), 195; and Hill and Blair, "Alien Oracles."

battlefield.[50] Whereas these aspects of knowledge all play a role in the targeting process, this discussion focuses on knowing-that and the ways in which this attribute can be ascribed to distributed cognitive systems.

Traditional epistemology describes knowledge as a justified, true belief. The application of this standard can be intuitive, though not without nuance. To have knowledge, one must first have a belief about some state of affairs. Further, that belief must also be true, meaning it must reflect an actual state of affairs. Finally, one must have adequate justification for that belief, which can be thought of as reasons that logically align with the belief itself. In short, one cannot know, for example, rain is falling without believing rain is falling and without rain actually falling. Even if one believes rain is falling, but has no reason (or has a bad reason) for that belief, one cannot claim to know rain is falling. For example, if one believes rain is falling because water is leaking from the roof, one likely has a reasonable justification for that belief. On the other hand, if one believes rain is falling because today is Tuesday (unless it frequently rains on Tuesdays), one likely has an inadequate justification for that belief.[51]

Empiricist epistemology views knowledge as originating in sense experience; thus, any justification for knowledge must be grounded in sense perceptions. Empiricists such as philosophers John Locke and David Hume argued the mind begins as a tabula rasa (blank slate), asserting mental perceptions, impressions, and sensations are unanalyzable givens that provide the raw data from which knowledge is built. In this view, abstractions such as concepts and beliefs must have clear connections to experience to have meaning or justification.[52] This view closely mirrors the data–information–knowledge–understanding paradigm described earlier: No knowledge or understanding can arise without data and information, which originate in sense perceptions.

Traditional concepts of knowledge tend to be truth preserving and prescriptive. These concepts are truth preserving because they aim to establish the truth of a matter, and they are prescriptive because they prescribe inferential processes

50. Audi, *Epistemology*, 225.

51. Robert M. Martin, *Epistemology: A Beginner's Guide* (Oneworld Publications, 2010), 5–15, 22–35; and Edmund L. Gettier, "Is Justified True Belief Knowledge?," *Analysis* 23, no. 6 (June 1963): 121–23. Gettier questioned the "justified true belief" model of knowledge by pointing out a belief can be true and justified but only accidentally count as knowledge. For example, suppose (1) Smith and Jones applied for a job, (2) Smith has strong evidence Jones will get the job, and (3) Smith also believes Jones has 10 coins in his pocket. Assuming Smith sees the entailment between the putative fact Jones will get the job and his having 10 coins in his pocket, Smith justifiably infers the person who gets the job has 10 coins in his pocket. But Smith does not know he has 10 coins in his pocket, and, as things turn out, he gets the job. Thus, Smith's belief the person who gets the job has 10 coins in his pocket is justified and true, but only accidentally. Whereas this concern has drawn a number of responses, it poses interesting implications for artificial intelligence, which, as addressed elsewhere in this study, frequently provides correct responses based on spurious correlations.

52. Martin, *Epistemology*, 107–14; and Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Harvard University Press, 1997), 11–17.

to guarantee truth. These conditions are often reflected in system design. For example, the Traffic Alert and Collision Avoidance System is truth preserving, as it must reliably detect the true positions and trajectories of nearby aircraft. To achieve such reliability, the system's sensors and algorithms are designed and tested against ground radar and aircraft transponders to ensure system-generated alerts correspond as closely as possible to actual airspace conditions. Once a collision risk is identified, the system prescribes a specific inferential process, providing pilots with complementary instructions to avoid a collision. For example, if one aircraft is told to descend, the other is told to climb.[53]

Although some uncertainty will always remain in practice, truth-preserving systems aim to reduce it as much as possible. To do so, especially in military contexts, systems tend to be hierarchical, with a clear chain of command and a division of labor in which personnel perform distinct roles. These systems are also deliberate, as roles are generally performed in sequence. For example, the targeting process begins with the collection of sensor data, followed by initial analysis and identification. The resulting identifications are then validated and further analyzed to provide an understanding of factors such as unit identity, mission, and readiness. Finally, commanders or their designated authorities decide which targets to strike with which assets.[54]

The logic employed in these systems tends to combine induction, which begins with observations and proceeds to general conclusions, and deduction, which involves reasoning from general conclusions to particular facts.[55] Inductive logic may be used initially to hypothesize a target based on observed behavior. For example, human operators in a tactical-level targeting process might observe a behavioral pattern: Whenever enemy combat vehicles move in a particular formation, an offensive operation is imminent. In this case, the logical form of inductive reasoning would adhere to the following sequence.

- Premise one: Every time enemy combat vehicles have been observed in a V formation, they have conducted an offensive operation.
- Premise two: These enemy combat vehicles are moving in a V formation.

53. W. H. Harman, "TCAS: A System for Preventing Midair Collisions," *The Lincoln Laboratory Journal* 2, no. 3 (1989): 437–58.

54. HQDA, *Army Targeting*, Field Manual 3-60 (HQDA, August 2023), 2-1–2-2; and O'Callaghan, "Fighting with the Data," 28–30.

55. Lorenzo Magnani, *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning* (Springer, 2009), 13–14.

- Conclusion: Therefore, these enemy vehicles are conducting an offensive operation.

These systems also reason deductively when identifying equipment or weapon systems in specific images. Here, deductive logic is used to ensure a target meets the operational parameters and rules of engagement. If one knows only supply trucks have three axles and travel exclusively in convoys, one can conclude a vehicle is a supply truck upon observing a three-axle truck traveling in a convoy. The logical form of deductive reasoning in this case would be as follows.

- Premise one: All supply trucks have three axles and travel in convoys.
- Premise two: These trucks have three axles and are traveling in a convoy.
- Conclusion: Therefore, these trucks are supply trucks.

Real-world reasoning is more complex. At the operational level, one may also draw on prior observations, noticing, for example, the enemy often places supply depots close to major roads or other critical transportation nodes. One may further observe supply depots are often located near heavily forested areas, where concealment is easier. At the strategic level, one may combine tactical and operational observations to generalize about when to prioritize supply depots over other targets, thereby informing the decision to strike. These facts can be combined inductively and deductively to strengthen the argument. As long as the observations are accurate and the logical rules are applied appropriately, the judgments based on them are sound.

So, suppose the preceding observations are true, and operators make the appropriate deductive and inductive connections among them. In this case, the operators can say they know the observed convoys consist of supply trucks, and their destination is a supply depot. But the fact a belief rests on sound reasoning does not entail certainty. As philosopher John Hawthorne points out, epistemic justification can take different modal forms, allowing one to treat probability claims as knowledge claims.[56] The Intelligence Community employs a similar concept when establishing the likely veracity of claims derived from analysis, as described in Intelligence Community Directive 203, *Analytic Standards*.[57] Thus, refining a knowledge proposition may include one's level of certainty. This accounting leaves space for knowledge regarding potential outcomes.[58]

56. John Hawthorne, *Knowledge and Lotteries* (Oxford University Press, 2004), 25–27.

57. Office of the Director of National Intelligence, *Analytic Standards*, Intelligence Community Directive 203 (Office of the Director of National Intelligence, June 2023).

58. Nicholas Rescher, *Epistemic Logic: A Survey of the Logic of Knowledge* (University of Pittsburgh Press, 2005), 49–50.

One can know, for example, a supply depot likely exists at a specific location. Discovering no such depot is present there does not change the veracity of the claim. Assuming appropriate reasoning, one could still claim the depot's presence was likely. This point illustrates the defeasibility of knowledge. Justifications may be momentarily acceptable, but they may be open to subsequent revision. This issue arises frequently in legal contexts. For example, if a witness claims person A shot person B, one may have a good reason to deduce person A actually shot person B. This reasoning could be defeated later if evidence emerged person A had an alibi.[59]

Nevertheless, this traditional approach to epistemology may not be as suitable for AI-enabled, distributed cognitive systems, especially those that rely on machine learning and neural networks. First, the traditional approach establishes a high standard, as successful knowledge claims depend on perfect information—in the sense of all observations being accurate—and flawless reasoning. This standard does not align well with how humans often reason, given imperfect senses and the complexities and uncertainties inherent in real-world situations. Second, AI-enabled systems operate in fundamentally different ways, using large amounts of data to uncover patterns and relationships that are neither explicitly defined nor easily interpretable by humans. Moreover, these systems learn as they are used. This learning allows AI-enabled components to adapt to changing situations and develop new capabilities. For example, AlphaGo, the first AI system to beat a human master, used reinforcement learning to discover novel strategies it had not encountered in human play.[60]

Additionally, the way AI-enabled systems create output, much less develop new capabilities, is somewhat inscrutable. Machine-learning systems using neural networks are often described as a black box in which AI-enabled processes are unobservable and, thus, unknowable. This inscrutability has multiple sources. First, many AI systems, especially those using deep neural networks, comprise layers of interconnected nodes that process large amounts of data in ways that are difficult to trace.[61] Second, unlike automated systems that employ software with explicitly programmed decision rules, machine learning results from statistical processes that are not explicitly coded, which can make the relationship between inputs and outputs less transparent.[62] These factors, among others, challenge the traditional, epistemic notion of knowledge as something fully understandable and explainable.

59. Douglas Walton, *Abductive Reasoning* (University of Alabama Press, 2005), 26–27.

60. David Silver et al., "Mastering the Game of Go Without Human Knowledge," *Nature* 550 (October 2017): 354.

61. Aggarwal, *Neural Networks*, 13–16; and John Zerilli et al., *A Citizen's Guide to Artificial Intelligence* (MIT Press, 2021), 70–73.

62. Aggarwal, *Neural Networks*, 16.

As noted earlier, the output of machine-learning systems is determined probabilistically. Because these systems generate predictions or classifications based on statistical inference rather than deterministic reasoning, their outputs express degrees of likelihood instead of certainties. The probabilistic character of machine-learning systems introduces a fundamental ambiguity about what knowledge one can draw from their use. The point is not simply about potential uncertainty regarding the truth of an AI-generated proposition. If this were the sole concern, one would only need to acquire more information to reduce uncertainty, as with induction. In fact, more information, even if it is relevant, can increase uncertainty by producing dilation, which is a widening of the range of possible probabilities rather than a narrowing of them.[63]

Instead, the difficulty stems from the opacity of the process AI-enabled systems use to justify their propositions. These propositions may follow from the statistical analysis of data considered by a system, but the specific data the system considered—or the ways in which these data were weighed—remain generally unknown. Thus, AI-generated representations of the world are not propositional truths; rather, they are distributions of confidence shaped by the quality, diversity, and representativeness of the data on which AI models were trained. Consequently, even when an AI system produces a correct answer, it may do so for reasons unrelated to the answer itself. For example, a well-known account describes a classifier that achieved high reliability in distinguishing between huskies and wolves. When the results were examined further, researchers found the classifier's high success rate was due to the presence of snow in the background of images of huskies but not wolves.[64] The algorithm excelled at classifying climate, not canines.

Applying truth-preserving standards requires clear, verifiable criteria for processes that are accessible—or at least comprehensible—to humans. Applying prescriptive standards requires modes of logic and inferential reasoning capable of resolving uncertainty while accounting for the strength of knowledge claims. In a sufficiently complex system, even one not enabled by AI, cognition is distributed both among human agents and, as will be discussed in the next section, among the tools and other artifacts the system employs. In such a system, humans often lack access to the range of criteria, processes, and reasoning the system employs to generate output.[65] So, because any particular human's cognition is less

63. Peter D. Grünwald and Joseph Y. Halpern, "When Ignorance Is Bliss," preprint, arXiv, October 25, 2005, https://arxiv.org/pdf/cs/0510080.

64. Chloé Andrews, "Why AI Fails in the Wild," Unbabel, November 15, 2019, https://unbabel.com/why-ai-fails-in-the-wild/.

65. Karin Knorr Cetina, *Epistemic Cultures: How the Sciences Make Knowledge* (Harvard University Press, 1999), 168–78.

than the whole and cannot access the whole (even collectively), no clear basis exists for grounding knowledge and, subsequently, trust.

Given the foregoing account, the epistemic challenge AI-enabled systems pose to mission-command processes should be easy to see. Whereas these systems can produce output that appears valid, they may justify their conclusions based on spurious correlations. Because these correlations are difficult to identify, especially in the heat of battle, AI-enabled systems can be fragile and fail as the operational environment diverges from the training dataset. This difficulty is compounded in GenAI and large language models, in which recommended conclusions and courses of action seem plausible but are ultimately untestable. As noted previously, output that suggests a judgment is likely does not become falsified simply because it turns out not to be true. So, a human- or machine-produced judgment, especially in assessing enemy intent or friendly courses of action, may not have worked out, but it may still have been the best judgment available given the circumstances. Acting on machine judgment, though resulting in apparent failure, may still have yielded the best outcome.

Still, treating a complex human-artifact system as a function of distributive cognition opens up new ways to understand the conditions for trusting AI-enabled systems. Consider an individual human mind. Humans trust their senses because the senses typically provide sufficiently reliable information about the environment. This reliability can be measured in terms of fit and function. Human senses often provide inaccurate information, and the environment contains many elements (such as segments of the electromagnetic spectrum) humans cannot directly perceive. Despite these limitations, humans have managed to survive and even thrive in various environments. This point suggests human cognition can achieve survival-related goals, despite not always being reliable.

Perhaps more to the point, the previous example of Grant at Vicksburg suggests humans may not always know, or be able to communicate, why a particular judgment is sound. This problem is also revealed in *Ender's Game*: Some AI judgments may rest on poor sensor input, bad data, or spurious correlations, but sometimes what appears to be bad output is actually advantageous. The question is: How would one know? Answering this question first requires an understanding of how knowledge is acquired in distributed systems, in which no single participating agent possesses the same knowledge as the entire system. Such understanding will entail system-design principles that allow the humans using systemic knowledge to trust the system, even when the justification is less than transparent.

Knowledge Production in Cognitively Distributed Systems

As Edwin Hutchins argues in *Cognition in the Wild*, cognition in distributed systems is not located within individuals' minds; rather, cognition is spread across people, tools, and environments in which the systems function. Through an ethnographic study of ship navigation, Hutchins described how the collective efforts of a ship's crew, coupled with the crew's navigational tools, create a complex cognitive system in which the computational properties of ship navigation do not reside solely in the heads of the sailors. In this system, crewmembers rely on both their individual knowledge and external representations (such as maps, instruments, and charts) to perform their tasks.[66] These tools extend sailors' cognition beyond the individual and the collective navigational system comprising humans and their tools.

To understand how cognition can be distributed in a system, one must first understand how elements of the system and the environment provide "cognitive scaffolding" that enables cognitive off-loading and burden sharing.[67] Cognitive scaffolding refers to the external structures—material, social, or technological—that support, extend, or enhance an agent's cognitive abilities. The concept originates in developmental psychology, in which scaffolding is described as the way teachers or parents structure tasks to help learners accomplish tasks they could not do alone. In cognitive science and philosophy, the concept has been expanded to include any environmental or artificial aid that helps an individual or a system think, remember, reason, or make decisions.[68]

In such systems, the cognitive properties of individuals, even when aggregated, are not the cognitive properties of the system as a whole. John Searle's famous thought experiment, known as the Chinese-room argument, illustrates this point (see figure 2). In the experiment, a man is placed in a room where he accepts input in the form of Chinese characters, consults a rule book that specifies a response, selects Chinese characters that match that response, and, finally, provides those characters as output. In this scenario, the man does not know Chinese, but, as Hutchins argues, the room does. Thus, knowledge of Chinese applies to the system, though not any part of the system, and certainly not the human in the room.[69]

66. Edwin Hutchins, *Cognition in the Wild* (MIT Press, 1995), 360.

67. "Cognitive Scaffolding," Fiveable, updated August 2025, https://fiveable.me/introduction-cognitive-science/key-terms/cognitive-scaffolding.

68. "Cognitive Scaffolding."

69. John R. Searle, *The Mystery of Consciousness* (*The New York Review of Books*, 1997), 11–12; and Hutchins, *Cognition in the Wild*, 362.

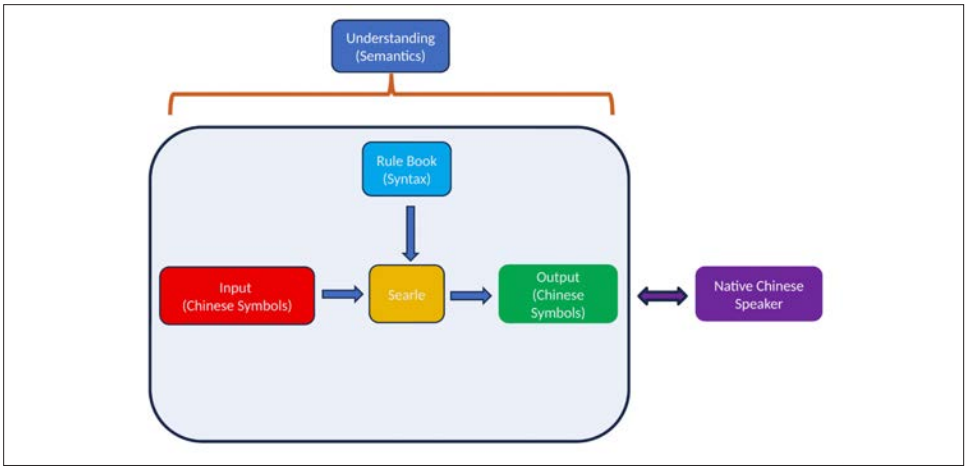

Figure 2. Cognitive map of the Chinese-room argument
(Source: Created by author)

At first glance, the idea of cognition extending beyond the mind, suggested by Hutchins's solution to the Chinese-room argument, seems odd, if not preposterous. But consciousness is not the same as cognition, and numerous subconscious processes—including memory retrieval, linguistic processes, and skill acquisition—are key to knowledge formation. Thus, the external nature of a process does not seem sufficient to disqualify the process from being cognitive.[70] In fact, biological brains frequently off-load cognitive tasks to the environment through the use of tools, written records, or other external aids. The physicist Richard Feynman, for example, observed his written work was not a record of his mental calculations; rather, the work was the calculations themselves.[71]

Ronald N. Giere and Barton Moffatt reinforce this point in their account of how humans off-load cognitive tasks to tools as simple as pencil and paper. Humans typically cannot multiply two three-digit numbers in their heads and must rely on pencil and paper to represent the steps required to arrive at the correct answer. In this cognitive system, the human contribution is to set up the problem, manipulate the representations in the correct order, and provide the products of any two integers, which can be done from memory.[72] More complicated systems, whether AI enabled or not, are capable of even more complex output. The difficulty, though, lies in more complex systems requiring more complex standards of trust before humans can depend on the knowledge the systems generate.

70. Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* (Oxford University Press, 2008), 222–24.

71. Clark, *Supersizing the Mind*, xxv.

72. Ronald N. Giere and Barton Moffatt, "Distributed Cognition: Where the Cognitive and the Social Merge," *Social Studies of Science* 33, no. 2 (April 2003): 303.

Differentiating the cognitive capabilities of systems and the individuals who compose those systems has several implications. The distribution of cognitive tasks enables the aggregation of expertise and reduces cognitive load. When processes are aligned with cognitive agents, they can lead to better and faster problem-solving and decision making. But such coupling also introduces challenges, including bias, misalignment, and complexities involving accountability and responsibility.[73] Because no individual in a system has the same cognition as the system as a whole, detecting deficiencies arising from these challenges—or knowing what to do, even upon detecting the deficiencies—may be difficult. Addressing these concerns through system and process design is critical to bolstering trust in a system.

But trust in a conventional, human-centric system is not simply a matter of attention to logical form and the proper alignment of a claim with its justification. Such trust also relies on confidence in the correct functioning of sensors and in the proper certification of the humans involved in operating the sensors. Trust also relies on the presence of human actors who are certified to interpret sensor output and able to oversee and validate the analytical process that transforms sensor data into information about enemy capabilities, knowledge of enemy intent, and wisdom to discern the right course of action. Because human cognitive capacities are limited and sometimes flawed, processes and certifications are insufficient for complete trust in the system.

Thus, as Karin Knorr Cetina observes, trust is also a function of epistemic culture, which distributes justification across an epistemic community comprising (much like Hutchins describes) humans, tools, processes, and institutions working in concert. Each component of this system comes with its own standards of trust. Whereas trust in individuals can be a function of professional expertise and the ability to deliver results, trust in the collective also depends on shared practices, methods, and norms. The trustworthiness of system-generated knowledge further depends on appropriate validation processes, social agreements, and institutionalized practices that ensure the credibility and reliability of the knowledge produced. In her study of high-energy physics experiments at the European research laboratory CERN, Knorr Cetina observed no single person, or even group of people, produced knowledge. Rather, the experiments themselves represented a kind of "collective consciousness" that served as an epistemic agent producing knowledge.[74]

73. Hollan et al., "Distributed Cognition," 176–77; and Arvind Narayanan and Sayash Kapoor, *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference* (Princeton University Press, 2024), 255–57.

74. Knorr Cetina, *Epistemic Cultures*, 23.

In Knorr Cetina's view, the Chinese room—not the human, who is just one component of the system—would be the "epistemic subject" capable of producing knowledge.[75] Trust, in this context, is a function of how well humans and tools collaborate. Trust in humans is a function of their experience, expertise, and reliability. Trust in tools is a function of their suitability and reliability. Process also matters because it specifies how collaboration occurs. Humans depend on shared practices, norms, and accountability to foster confidence in their collaboration. Tools require processes of calibration, validation, and standardization. At the experimental level, an institutional capacity must exist to review, replicate, and disseminate results.[76]

As Giere and Moffatt also argue, ascribing cognitive agency to cognitively distributed systems, as Hutchins and Knorr Cetina do, is tempting, as it allows one to say the system knows, perhaps even consciously. But Giere and Moffatt argue one should resist that temptation. In their account, systems can be epistemic agents (that is, producers of knowledge) without being cognitive agents (that is, holders of knowledge). Instead, only humans can claim knowledge of the system's output. In fact, one need not be a part of the system to derive knowledge from it. Whereas humans may come to know the output as part of the cognitive system, most learn about that output through authenticated reports. Thus, the act of knowing is essentially the same for both human participants and observers—only the justification for knowledge is different.[77]

Giere and Moffatt's approach is intuitively appealing. Accepting the idea that tools allow human cognition to extend beyond the mind is one thing; accepting a cognitively distributed system as a mind is another, even if the system and the mind share similar functions. By differentiating knowledge production from knowledge acquisition, Giere and Moffatt's view narrows any knowledge gap between what a cognitively distributed system produces and what humans can say they know. Trust in the system's output depends on the quality of human participants' expertise and experience, properly calibrated tools, and effective processes.

But even though this view may narrow one knowledge gap, it risks expanding another. Legacy systems capable of producing knowledge also rely on relationships among humans, tools, and processes that are clearly defined and applied within a limited scope of information, as evidenced by legacy military decision-making and

75. Knorr Cetina, *Epistemic Cultures*, 178.

76. Knorr Cetina, *Epistemic Cultures*, 133–36.

77. Giere and Moffatt, "Distributed Cognition," 304.

targeting systems.[78] Moreover, whereas aspects of the system's operation may not be available to individual humans, they are generally understandable. The person in the Chinese room may not know what the individual Chinese characters mean, but that person knows how the rule book works.

This comparison between probabilistic and intuitive reasoning underscores a deeper epistemic challenge in mission command: how to generate and act on knowledge in the face of uncertainty.[79] Whether derived from human intuition or machine inference, both forms of judgment depend on integrating incomplete, context-dependent information into a coherent operational picture that provides commanders with the understanding they need to make sound decisions. The issue, then, is less about whether AI can replicate human reasoning and more about how human and machine cognition can jointly sustain reliable sensemaking. Addressing these trust challenges requires, first, understanding the system itself; second, examining how the system produces knowledge claims; and, third, determining how those claims can be assessed for epistemic reliability. Together, these conditions establish the foundation for viewing battle command as a cognitively distributed process in which knowledge, judgment, and decision authority are shared across humans, machines, and the operational environment.

Cognitive Distribution, Mission Command, and the Targeting Process

As Richard Hart Sinnreich points out in an early study of battlefield cognition, battle command, which for the purposes of this discussion is synonymous with mission command, is the "art and science" of decision making that leads to battlefield success.[80] In this view, mission command poses both puzzles and mysteries, constituting a creative yet difficult-to-quantify process—one constrained by predefined objectives that require confirmable principles whose application yields predictable results. Cognitive tasks associated with effective battle command include diagnosing, planning, deciding, synchronizing, and communicating. Diagnosing requires commanders to infer patterns from incomplete information to assess the potential to achieve battlefield objectives. Planning, given this irreducible uncertainty, is best understood as organizing resources to gain an advantage at the

78. Gary Kuczynski, *Military Decision-Making Process: Organizing and Conducting Planning* (Center for Army Lessons Learned, 2023); and HQDA, *Army Targeting*, 2-15. Both sources describe hierarchical, top-down knowledge-production and decision-making systems.

79. Sinnreich, "Variables and Constants."

80. Sinnreich, "Variables and Constants," 12. In the context of the present discussion, battle command refers to the policies, processes, people, tools, and other components necessary for controlling forces in combat. Mission command is a particular arrangement of these components.

outset of battle and retain that advantage as the battle evolves. Deciding involves translating solutions from the planning process into intentions on which others act. Acting on these intentions requires synchronizing them in time, space, and purpose to maximize effects. It also requires effective communication and motivation as the battle unfolds.[81]

Carrying out these tasks successfully depends on good situational awareness. In the same study of battlefield cognition, Micah R. Endsley offers a helpful characterization of the three levels of situational awareness, highlighting the epistemic requirements of successful battle command. At the first level, agents perceive information about the nature and location of an object or objects of interest (for example, an enemy tank). At the second level, agents comprehend the meaning of the object or objects to understand what the objects are doing and could do in the future. At the third level, agents can integrate relevant environmental context to project the situation over time and reasonably approximate how it might evolve.[82]

By drawing a map of the cognitive processes of battle command and targeting, one can provide a clearer sense of the epistemic requirements for success, making explicit how information is transformed into understanding at each stage. By showing how the levels of situational awareness—perception, comprehension, and projection—interact with the tasks of diagnosing, planning, deciding, synchronizing, and communicating, the map reveals what kinds of knowledge processes are required (for example, recognizing patterns, anticipating enemy actions, and aligning intentions with resources) and where gaps or failures in those processes could occur. The map also shows command is iterative and distributed: Knowledge emerges from the interplay of people, tools, and context rather than from a single decisionmaker. This mapping helps analysts and practitioners identify where epistemic challenges—such as uncertainty, bias, or overload—are most likely to affect outcomes, and where interventions—such as enhanced training, improved information systems, or AI tools—are most needed.

Robert Axelrod, in an early text on cognitive mapping, proposed the basic elements of a cognitive map are the concepts people have, the causal links between those concepts, and the alternatives among which decisionmakers must choose.[83] For Axelrod, cognitive maps consist of concept variables, represented by points, and causal connections between those variables, represented by arrows. A causal connection can be either positive, where an increase in one variable leads

81. Sinnreich, "Variables and Constants."

82. Micah R. Endsley, "Situational Awareness: A Key Cognitive Factor in Effectiveness of Battle Command," in Kott, *Battle of Cognition*; and Gary L. Klein et al., "Enabling Collaboration: Realizing the Collaborative Potential of Network-Enabled Command," in Kott, *Battle of Cognition*.

83. Robert Axelrod, ed., *Structure of Decision: The Cognitive Maps of Political Elites* (Princeton University Press, 1976), 5.

to an increase in the other, or negative, where an increase in one variable leads to a decrease in the other. A sequence of distinct points connected by arrows is a "path," which can also form cycles in which the variable at the last point affects the variable at the first point.[84]

But Axelrod was concerned only with individual cognition. To understand distributed cognition, one must first map the system or process doing the cognizing. For the sake of simplicity, the following discussion focuses on kinetic targeting that requires a system capable of deciding what, when, where, and how by detecting and identifying targets, delivering appropriate capabilities that create the desired effects, and assessing those effects to inform future operational decisions.[85]

To draw a cognitive map of the targeting process, the discussion begins with the D3A cycle as an example of a battle-command process. Then, the discussion charts the flow of cognition by linking each map element to the situational-awareness levels that enable it: perception of objects and events (first level), comprehension of their meaning and potential actions (second level), and projection of how the situation may evolve (third level). Arrows between nodes indicate the direction of influence or dependency—for example, how diagnosing patterns based on incomplete information feeds into planning, or how planning shapes decisions that must then be synchronized and communicated. To capture complexity, feedback loops show how later tasks, such as synchronization, can inform earlier tasks, such as diagnosis, creating an iterative cycle rather than a linear chain. The resulting map is both a visual representation of distributed cognition in targeting and a diagnostic tool for identifying where uncertainty, information gaps, or breakdowns in situational awareness may undermine effective battle command.

Army Targeting, Field Manual 3-60, lists 26 different roles agents perform in the D3A targeting process. To simplify, this discussion considers only the roles aligned with the functions specified in the cycle. For present purposes, these roles include the commander, the intelligence officer (G-2), the operations officer (G-3), the fire support coordinator (FSCOORD), the plans officer (G-5), the assessment working group, and the staff judge advocate.[86] In the context of battle command, drawing on James D. Thompson's 1967 study, *Organizations in Action*, Gary L. Klein et al. argue three kinds of processes link these agents. The first kind is a long-linked process in which agents accomplish tasks over time and in sequence. An intelligence process that links collection, surveillance, and analysis would be one such process. The second process is a mediating link that connects agents to each other, enabling them to gather information and coordinate action, much as a broker

84. Axelrod, *Structure of Decision*, 72.

85. HQDA, *Army Targeting*, 2-1–2-2.

86. HQDA, *Army Targeting*, 1-8–1-14, 2-13.

mediates between buyers and sellers of stock. In the context of targeting, effects management involves a mediative process of connecting targets with appropriate assets, given each target's characteristics. The third process is intensive, focusing on making changes in the environment and relying on information about prior changes to inform future ones.[87]

These agents are also related through pooled, sequential, and reciprocal interdependence. In pooled interdependence, agents provide discrete contributions that are pooled to create a more comprehensive understanding. For example, a specific targeting process may draw on separate streams of signals, geospatial, and human intelligence, which are then fused into a single estimate. Sequential interdependence means the output of one agent depends on another agent's output. The D3A process itself is an example of such interdependence, as the delivery part of the cycle depends on fulfilling the parts pertaining to decision and detection. Reciprocal interdependence means agents (or their organizations) depend on the actions of others before acting themselves. One example of such interdependence is the relationship between fires and sustainment: Logistics must first deliver munitions and maintenance to enable repeated strikes, and the effects of those strikes and their contribution to operational maneuver shape resupply and maintenance requirements and priorities.

The decision phase initiates the targeting cycle by identifying which targets must be engaged to achieve the commander's operational objectives. This process is informed by the commander's guidance, the scheme of maneuver, and the anticipated enemy threat. The commander plays a central role in the process, issuing intent and objectives that guide the rest of the staff. The G-2 contributes threat analysis and leads the intelligence preparation of the battlespace, whereas the G-3 ensures synchronization with maneuver planning. The FSCOORD chairs the targeting working group, managing the development of the high-payoff target list, the attack-guidance matrix, and the target-selection standards—the essential products of this phase. The staff judge advocate ensures legal and ethical compliance. Interactions at this stage are collaborative, structured around the targeting working group, and foundational for the phases that follow.[88]

In the detection phase, the focus shifts to locating and confirming the targets identified during the decision phase. The G-2 and the collection management team lead the effort to task and synchronize intelligence, surveillance, and reconnaissance (ISR) assets in accordance with the commander's priority intelligence requirements. The G-3 supports task execution by managing operational timelines and ensuring sensor tasking aligns with the overall operation. The FSCOORD and the targeting

87. Klein et al., "Enabling Collaboration."

88. HQDA, *Army Targeting*, 2-2–2-5.

officer help ensure detection efforts are tied to actionable targeting decisions, keeping the sensor-to-shooter pipeline coherent. Interactions are time sensitive and data driven, often requiring the integration of national-level and organic ISR assets. Intelligence and operations staff must collaborate closely to validate, refine, or reprioritize targets based on detection results, setting conditions for timely and effective engagement.[89]

The delivery phase involves engaging validated targets through lethal or nonlethal means. Based on the attack-guidance matrix and fire support planning, the G-3 directs the employment of available capabilities, whereas the FSCOORD manages targeting procedures and deconfliction across Joint fires. The commander retains ultimate authority over target engagement, though this authority may be delegated depending on the rules of engagement and the operational tempo. Asset managers assess whether the designated effect can be achieved with available weapons or nonlethal tools (for example, cyberspace operations, electromagnetic warfare, and information operations). The coordination between operations, fires, and intelligence is essential to ensure engagements are lawful, discriminate, and synchronized with the larger scheme of maneuver. The success of this phase hinges on accurate detection, rapid data transfer, and well-rehearsed command relationships.[90]

The assessment phase closes the targeting loop by evaluating whether the intended effects were achieved and by determining whether further action is needed. This phase begins with battle damage assessment, which is led by the G-2 and includes physical, functional, and systemic evaluations of the target and its context. The G-3 and fires staff provide operational input and may assess the effectiveness of munitions, whereas the G-5 or the designated assessment working group consolidates findings to inform future operations. Feedback from the assessment phase directly supports adjustments in the decision phase, creating a continuous cycle of refinement. Interactions across intelligence, operations, and fires ensure assessments are comprehensive and grounded in both tactical effects and strategic impacts, reinforcing the adaptive and iterative nature of the D3A framework.[91]

Targeting processes do not take place in isolation. Instead, they are part of an ecosystem in which at least two adversarial agents target each other. In this interaction, the strategic problem for each agent is to optimize action against the

89. HQDA, *Army Targeting*, 2-6–2-8.

90. HQDA, *Army Targeting*, 2-7–2-10.

91. HQDA, *Army Targeting*, 2-11–2-14.

adversary's behavior without knowing what the adversary will do.[92] In this ecosystem, speed matters: The advantage goes to whichever party gets through the targeting cycle faster. Targets, in this context, are grouped into two categories: deliberate and dynamic. Agents develop deliberate targets over time, whereas dynamic targets are unanticipated opportunities. Strike assets are apportioned to deliberate targets as part of the targeting process. Dynamic targets lack clear priorities or dedicated assets and, most importantly, the necessary coordination to deliver a strike rapidly. This fact adds a layer of complexity to the process of sorting dynamic targets, which can take time and can allow the enemy to escape out of range, find cover, or otherwise reduce its vulnerability.[93]

Distributed cognition produces knowledge in the targeting ecosystem by connecting specialized agents, information flows, and procedural artifacts in a coordinated, knowledge-generating process. In this process, individual roles—such as the commander, the G-2, the G-3, the FSCOORD, the G-5, the staff judge advocate, and the assessment working group—contribute discrete observations and judgments that are pooled, passed sequentially, or exchanged reciprocally across long-linked, mediating, and intensive processes. Shared products—such as the high-payoff target list, the attack-guidance matrix, the target-selection standards, the priority intelligence requirements, and time-sensitive sensor tasking—transform raw perception into validated hypotheses that are applied to doctrinal checklists, legal review, and commander intent, providing the prescriptive scaffolding that turns hypotheses into deductively justified actions. As the cycle iterates, assessment feeds back into the decision phase, allowing priorities to be revised and beliefs to be updated.

Knowledge, therefore, does not reside in a single mind or sensor; rather, knowledge emerges from the integration of heterogeneous inputs, role-specific processing, and repeated D3A cycles—with speed, corroboration, and the ability to reconfigure around dynamic targets determining epistemic superiority in an adversarial setting. Given this complexity, providing a complete map of the cognitive distribution in this system is well beyond the scope of the present discussion. Figure 3 shows a simplified version of the map, illustrating the non-AI-augmented system.

92. Susan E. Riechert and Peter Hammerstein, "Game Theory in the Ecological Context," *Annual Review of Ecology, Evolution, and Systematics* 14 (1983): 382.

93. HQDA, *Army Targeting*, 2-8.

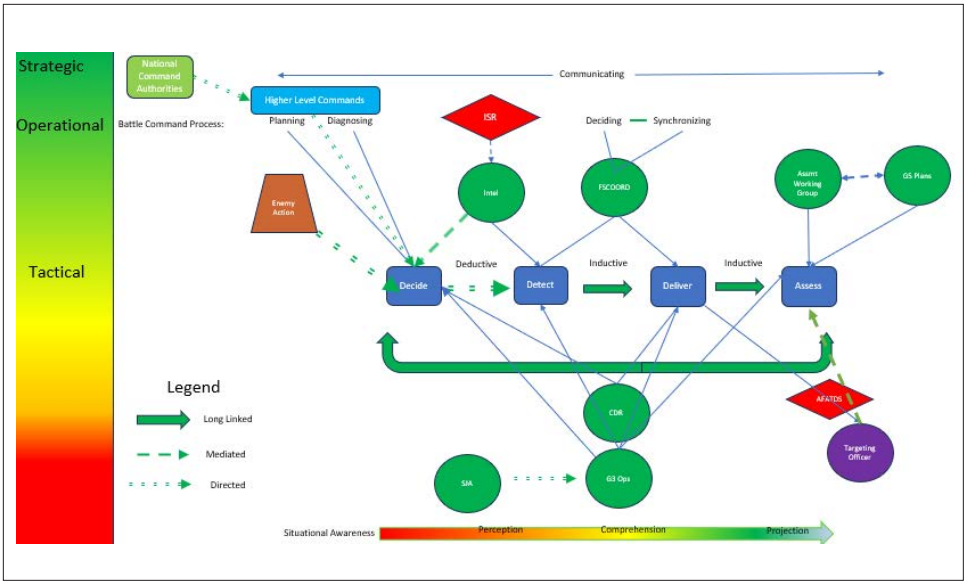

Figure 3. Cognitive map of the legacy battle-command and targeting process
(Source: Created by author)

Epistemology and Design

To understand how AI augmentation impacts the distribution of cognition, one must first understand how design affects knowledge production. Understanding the targeting process as a cognitively distributed system enables one to describe how that system produces knowledge. But such understanding does not provide an account of whether the system produces knowledge. To attain this account, one must know two things: (1) whether the design is appropriate for its stated purpose, and (2) on what basis the design's output can be trusted. At first glance, these standards may seem like a distinction without a difference. If the design is appropriate for the task, its output can be trusted. But an organism's design being well suited to an environment does not mean the organism's cognition cannot be flawed under certain circumstances. This point is aptly illustrated by the fact one's senses occasionally provide false or incomplete information about one's environment.[94]

Regarding the first standard, the black-box design of AI models and the probabilistic nature of their outputs complicate both the process of establishing

94. One may argue organisms are products of random evolution and are not designed; therefore, the analogy breaks down. But an organism not being designed does not mean it does not have a design whose appropriateness can be measured against the standard of survival in a particular environment. As discussed later, many organisms alter themselves and their environment—that is, alter their design or improve their fit—to reduce selective pressures.

whether AI output is accurate and the process of determining whether the output is more effective than the one obtained via human processes. As Federica Russo et al. point out, the epistemology of AI should account for the reliability and precision of AI models or, if such accounting is not possible, establish the conditions under which one can trust the models' results.[95] According to philosopher Rico Hauswald, expertise and epistemic authority are different. One can be an expert, but unless one can transfer that expertise to a nonexpert, one is not an epistemic authority.[96] The transfer requirement thus adds another dimension to the epistemic role AI plays as a knowledge producer.

As mentioned earlier, legacy systems are fundamentally empiricist in design and operation, assuming reliable knowledge about the battlespace can be derived from systematic observation, measurement, and correlation. For example, activities such as intelligence collection, sensor fusion, and battle damage assessment rely on an empiricist model in which agents gather sensory data, verify the data's accuracy, generalize patterns, and then act. This epistemic cycle relies on inductive reasoning, in which knowledge emerges from accumulated observations. For illustration, suppose a commander has decided to locate and destroy an enemy's indirect-fire capabilities and sends forces to the area where these capabilities are believed to be positioned. During the movement, a forward observer might observe muzzle flashes and report them back to the G-2, who may then direct other ISR assets, signals, or other resources to collect additional information.

Whatever the additional observations are, they are combined inductively to form the hypothesis an enemy artillery battery is active at a specific location. Once the evidence reaches a set confidence threshold, the commander deduces a course of action, such as authorizing a strike on the location. This reasoning is deductive because the commander is moving from accepted premises—such as the inductively derived location of the artillery battery, a doctrinal rule according to which the disruption of enemy artillery fires is necessary to maintain freedom of maneuver, and operational assumptions regarding what actions are required to disrupt enemy artillery—to ordering a strike. After the strike, additional observations confirm or disconfirm the hypothesis, completing the empirical feedback loop.

Described this way, legacy targeting processes build knowledge based on repeated sense perceptions, raise confidence by corroboration, use prescriptive rules to justify action, and rely on the deductive application of justified premises to make decisions and act. Throughout these steps, the legacy targeting process

95. Federica Russo et al., "Connecting Ethics and Epistemology of AI," *AI & Society* 39 (2024): 1588.

96. Rico Hauswald, "AI and the Philosophy of Expertise and Epistemic Authority," in *A Companion to Applied Philosophy of AI*, ed. Martin Hähnel and Regina Müller (Wiley, 2025), 63–66.

remains vulnerable to the problem of induction—namely, the fact induction can support high confidence but never certainty. Addressing this fundamental problem depends heavily on perceptual fidelity, in which multiple observations from various sensors and sources provide the optimal resolution of the perceived world. Addressing uncertainty also depends on effective rules for acting—for example, the rules prescribed in military doctrine—when perceptions cross certain thresholds, as the earlier example illustrated.

The epistemic challenge lies primarily in the limitations of sensory capabilities and human cognition, particularly in perceiving, interpreting, and reasoning about a complex environment. In such a system, knowledge is justified by shared norms of evidence, deliberation, and expert judgment, with reasoning being transparent and corrigible through dialogue or review processes that are integrated into the system. Systems and processes are thus designed with these norms in mind, allowing knowledge claims to be traced to observable evidence, evaluated through transparent and reproducible procedures, and corrected should new evidence or rules arise.

Epistemology and AI Design

In contrast to legacy systems, AI-enabled systems introduce a fundamentally different epistemic problem: establishing the limits of intelligibility rather than explainability. Addressing this problem requires different solutions. In AI-enabled systems, knowledge is not simply interpreted by humans; rather, knowledge is coproduced with opaque, probabilistic models whose internal logic often resists explanation in human terms. The resulting epistemic challenge concerns both whether beliefs (and the decisions they support) are accurate or valid and whether the system provides the right kind of cognitive scaffolding to transform data into understanding.

As discussed, AI-enabled systems reshape how knowledge, problem-solving, and decision making are distributed across individuals and systems. First, these AI-enabled systems amplify cognitive capacity by rapidly processing vast amounts of data and identifying patterns that would exceed human capabilities, even without time or personnel constraints. Second, these systems automate routine cognitive tasks such as reviewing sensor feeds, detecting and classifying targets, and prioritizing engagements.[97] Automation frees human operators to focus on higher-order cognitive functions such as interpretation, judgment, and intent

97. C. Anthony Pfaff and Christopher John Hickey, *Integrating Artificial Intelligence and Machine Learning Technologies into Common Operating Picture and Course of Action Development* (US Army War College Press, 2025), 7–9, https://press.armywarcollege.edu/monographs/980/; and O'Callaghan, "Fighting with the Data," 21.

formation. As GenAI capabilities are integrated into AI-enabled systems, machines may begin performing aspects of these more abstract tasks as well. For instance, machines may go beyond simply recognizing a formation of combat vehicles, further inferring those vehicles' likely role or purpose in the current tactical context. Third, by assuming advanced analytic and interpretive roles, AI-enabled systems can facilitate more effective cognitive off-loading, allowing human-machine teams to operate more efficiently and coherently across larger scales of data, time, and complexity (see figure 4).

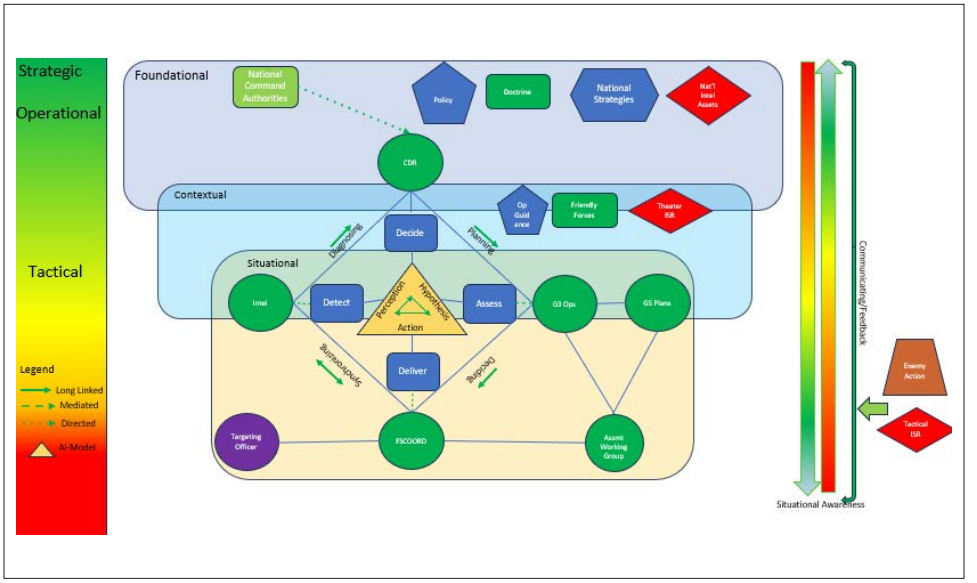

Figure 4. Cognitive map of an AI-enabled D3A cycle
(Source: Created by author)

Seeing how AI-enabled systems change the traditional understanding of war fighting is not difficult. In such systems, humans can operate at higher levels of cognition, think more abstractly, and be more responsive to new information. In the context of targeting, as in nature, using these systems should enhance the effectiveness and, consequently, the survival of a system and the humans who participate in it. But the question remains: What epistemic standards are appropriate? As mentioned earlier, targeting systems operate in conjunction with allied and adversarial systems. These joint operations occur within a targeting ecosystem, in which competing and complementary systems interact in a cycle that iterates with other systems.

Despite its advantages, increased cognitive off-loading enabled by AI also introduces an epistemic risk that is not present in legacy systems. As more cognitive work is delegated to machines, human operators may lose insight into how conclusions are reached, eroding their ability to detect when system outputs

deviate from context or intent. Over time, these effects can produce a false sense of confidence in machine-generated insights, with efficiency being gained at the expense of understanding—precisely the condition most likely to undermine trust and decision quality in complex, adaptive environments.

Efforts to address this concern must ensure AI-enabled systems are designed to provide reliable output while mitigating environmental pressures that threaten survival and impede goal attainment. One approach designers rely on to validate AI output is "computational reliabilism," which measures reliability in terms of indicators that, when present, justify believing a system's output is accurate.[98] These indicators include (1) verification and validation measures, (2) robustness, (3) history of success and failure in implementation, and (4) expert knowledge. The first two indicators address the technical aspects of the system, and the last two indicators address the context in which the system operates.[99]

Verification and validation refer to the correctness of the model employed by the system and the model's ability to yield accurate results. The second indicator, robustness, refers to the model's ability to emulate the real world across a variety of circumstances. The third indicator refers to the history of the system's successes and failures, serving as a cumulative indicator that captures when the system succeeded or failed and the circumstances under which these outcomes occurred. The fourth indicator, expert knowledge, plays a role in the evaluation of the other indicators. Expert knowledge requires humans to assess whether the assumptions underlying the model's design are adequate to determine the model's correctness, validity, and robustness.[100]

Like the empiricist epistemology described earlier, computational reliabilism draws its justification from observations of system performance and inductive generalizations based on these observations. Also like empiricism, computational reliabilism assumes the consistent production of accurate results is sufficient to justify belief in a system's reliability. But this assumption is inadequate for an AI design epistemology because it treats reliability as a statistical property of outputs rather than as an epistemic property of the processes that generate those outputs. In complex environments, a system's past performance offers little assurance the system will continue to perform reliably in novel situations. More importantly, such an approach does not assess the utility of system output, as something can be true without being useful. So, rather than reliability, the important factor is understanding the niche into which the system fits in the operational environment and how that fit enables advantages over competitors.

98. Russo et al., "Connecting Ethics and Epistemology," 1589–90.

99. Russo et al., "Connecting Ethics and Epistemology," 1589–90.

100. Russo et al., "Connecting Ethics and Epistemology," 1589–90.

In place of an empiricist approach, Attila Márton Putnoki and Tamás Orosz argue for a metric they call "cognitive resonance," defined as "the degree of correspondence between system recommendations and the evolving preferences and expectations of the decision-maker, conceptually ranging from 0 (no alignment) to 1 (perfect alignment)."[101] As the authors observe, conventional AI models focus on "task-specific optimization of outputs" and are less useful in more complex decision making that involves human agents making subjective assessments.[102] Instead, complex decision making requires a system that aligns with a human user's cognitive structures, reasoning practices, and subjective preferences—one able to generate hypotheses and adapt based on feedback from acting on those hypotheses.[103] The authors propose a six-domain "Cognitive and Artificial Intelligence Evaluation" model for aligning machine processes with human cognition. These six domains are "Learning and Adaptation," "Perception and Awareness," "Reasoning and Decision-Making," "Interaction and Personalization," "Memory and Retention," and "Optimization and Operational Efficiency."[104]

In this framework, learning and adaptation reflect a system's ability to acquire new knowledge and extract new patterns as the operational environment evolves. Perception and awareness refer to the system's ability to process data from multiple inputs, creating situational awareness that supports anomaly detection. Reasoning and decision making reflect the logic of problem-solving, the system's handling of trade-offs, and the alignment of decisions with goals. Interaction and personalization refer to the system's ability to adapt to both the environment and the user, enabling better decision making. Memory and retention refer to the management of short- and long-term memory, which is necessary to build on acquired knowledge. Optimization and operational efficiency refer to the effective management of computational resources to avoid unsustainable costs or unexpected vulnerabilities.[105]

Understanding mission command as distributed work suggests an important point: Commanders and staffs do not need total explainability to use AI-enabled systems effectively. As mentioned earlier, uncertainty in these systems is unresolvable. But the preceding discussion also clearly indicates this challenge is often present in mission command without AI-enabled systems. In practice, commanders and staffs often use tools they do not fully understand internally. Human trust in those tools exists because they are calibrated, bounded by processes, and embedded

101. Attila Márton Putnoki and Tamás Orosz, "Evaluation and Adaptive Complexity of Cognitive Information Systems," *Acta Technica Jaurinensis* 18, no. 4 (2025): 227.

102. Putnoki and Orosz, "Evaluation and Adaptive Complexity," 225.

103. Putnoki and Orosz, "Evaluation and Adaptive Complexity," 225–26.

104. Putnoki and Orosz, "Evaluation and Adaptive Complexity," 225–28.

105. Putnoki and Orosz, "Evaluation and Adaptive Complexity," 227–28.

in a wider professional culture of checking and challenge. The problem with some AI systems goes beyond their opacity; these systems can be opaque while also changing the basis on which commanders and staffs attend, prioritize, and interpret. Rather than merely providing an output, these systems shape the conditions under which commanders and staffs form beliefs. So, where conditions for belief change, so too should the way one forms beliefs.

Ecocognitive Epistemology and AI

The idea of cognitive resonance fits well into an ecocognitive approach to knowledge and understanding. The domains proposed by Putnoki and Orosz correspond to the cognitive functions agents need to fit into their environment. Sustaining that fit depends on three subordinate functions: representation, hypothesis formation, and feedback. The ability to represent environmental features adequately is necessary to support reasoning and decision making. Reasoning identifies the relevant environmental features required for hypothesis formation and evaluation. Decisions based on that evaluation lead to actions that can affect both the agent and the environment, including how the agent interacts with—and sometimes changes—the environment to achieve its goals. Assessing these interactions requires an effective feedback mechanism that allows the agent to acquire new information and adapt to its environment.

An ecocognitive approach also nests well within how humans interact with the world around them. In *Naturalizing the Mind*, epistemologist Fred Dretske argues mental phenomena should be understood as part of the natural world, governed by the same laws as physical processes. In Dretske's view, concepts such as perception, intentionality, and knowledge can be explained in terms of how they fit into processing information about the natural world.[106] This shift in focus, argues Hilary Kornblith, differentiates abstract concepts of knowledge from the natural phenomenon of knowledge as experienced by humans.[107] This point does not suggest abstractions play no role; rather, it suggests the role of abstractions is determined by how well the abstractions fit within the system, its goals, and the environment.

By grounding epistemology in the natural world, knowledge and understanding enable agents to process and respond to environmental stimuli.[108] In this context, knowledge is goal oriented, and trust-related standards should be goal oriented as well. Being goal-oriented agents, living organisms cannot separate perception from

106. Fred Dretske, *Naturalizing the Mind* (MIT Press, 1997), 72, 162–68.

107. Hilary Kornblith, "Naturalistic Epistemology and Its Critics," *Philosophical Topics* 23, no. 1 (Spring 1995): 243–44.

108. Bertolotti, *Patterns of Rationality*, 5.

problem-solving. Moreover, rather than passively receiving information from the environment, these agents can produce effects based on their inferences, potentially altering the environment in ways that are either advantageous or disadvantageous to themselves.

Unlike the empiricist view described earlier, ecocognitive epistemology is neither truth preserving nor broadly prescriptive. This form of epistemology is not truth preserving because it relies on abduction as its principal means of inference. Unlike deduction and induction, abduction is not truth preserving because it employs a standard of plausibility rather than certainty, and abduction is not prescriptive because it aims to describe the available explanations, with an emphasis on avoiding errors when choosing the best explanation.[109] Drawing on Kay and King's distinction between puzzles and mysteries, abduction is the logic of mystery solving. It is akin to the reasoning the fictional detective Sherlock Holmes employs when solving a case, and abduction is especially useful when the facts at hand underdetermine the solution.[110] In such cases, the available facts can combine in many different ways to provide a solution, but determining which explanation is best remains impossible. One can only assess which explanations work. In Holmes's case, radical uncertainty is resolved only after the fact, only if the suspect confesses. In the case of mission command, radical uncertainty is also resolved only after the fact, when one can determine whether the act premised on the abduction achieved the intended goal.

What commanders want, of course, is a way of assessing whether a particular explanation is sufficiently good to justify action. Here, ecocognitive epistemology can help not so much by distinguishing better and worse explanations, but rather by assessing whether the mission-command system fits the operational environment well enough to generate good-enough explanations. Ecocognitive epistemology emphasizes the interaction between cognition and the environment in the process of knowledge acquisition. In this view, knowledge arises both from internal cognitive processes and from external environmental factors, shaped by the tools or technologies humans use to interact with the world. Therefore, any standards for epistemic success should account for how interactions between cognition and the environment fit the adaptive nature of cognition in real-world contexts, such as in military targeting systems. From the perspective of ecocognitive epistemology, knowledge of the system's appropriate design, the reliability of its output, and the alignment of its behavior with its goals are functions of how the system fits into the environment.

109. Bertolotti, *Patterns of Rationality*, 5–14.

110. Kay and King, *Radical Uncertainty*, 224–25.

Artificial Intelligence and Directions of Fit

Fit, in the context of ecocognitive epistemology, is relative. A mosquito, for example, has much more limited cognitive capabilities than a human, yet it remains well adapted to its environment: The mosquito can sense what it needs to survive, and it can detect threats to that survival. To sustain its place in the ecosystem, a mosquito compensates for its vulnerabilities by rapidly reproducing in large numbers. The complexity of cognition is less important than the appropriateness of the relationship between the organism and its environment. The same holds true for AI-enabled cognitive systems. These systems' effectiveness depends less on computational power or data access than on the alignment of the systems' perceptual, inferential, and adaptive functions with the operational environment and the human cognition with which they interact. The fit between cognition and the environment refers to the correspondence between internal representations and the external world that enables effective action. But this fit is not unidirectional. Whereas environmental factors place selective pressure on agents, agents often respond to that pressure by modifying the environment. Thus, assessing fit depends first on understanding the fit's direction.

In the context of cognition, Searle argues the direction of fit explains how the mind relates to the world and optimizes action. For example, beliefs have a "mind-to-world" direction of fit, meaning they should fit how things are in the world.[111] If a belief does not satisfy that condition, one should change the belief. Desires and intentions, on the other hand, have a "world-to-mind" direction of fit, in which the aim is to make the world conform to one's mental state; failing that conformity, one should either change the world or abandon the desire or intention.[112] Searle uses this distinction to explain practical rationality, a function critical to problem-solving. In this context, acting effectively involves a "double direction of fit" in which an intention to act represents both how the world is and how it should be.[113]

Searle's directions of fit map neatly onto the cognitive architecture of AI-enabled targeting systems: The mind-to-world direction associated with belief corresponds with those systems' cognitive functions, especially the functions of diagnosing and assessing who is responsible for representing the battlespace as it is. For example, an AI model that fuses ISR, human intelligence, and signals intelligence might produce a probability estimate that suggests a hostile artillery battery occupies a particular location, thereby operating with a mind-to-world orientation. The model is justified in its output insofar as it produces representations that correspond with

111. John R. Searle, *Rationality in Action* (MIT Press, 2001), 39.

112. Searle, *Rationality in Action*, 39.

113. Searle, *Rationality in Action*, 36–39.

reality—in this case, the presence of an artillery battery at the identified location. By contrast, the world-to-mind fit associated with desire and intention corresponds with other cognitive tasks—such as planning, deciding, and synchronizing—that aim to make the world conform to the system's goals, which presumably involve destroying or neutralizing the enemy artillery battery. The desires and intentions behind the system's goals are justified to the extent the system can realize them.

The double direction of fit that characterizes the cognitive processes of AI-enabled targeting systems means effective action requires representations that both accurately describe the world and are sensitive to the normative constraints of an intended change. In this framework, normativity is relative to a system's functions: If the system is designed to do something—and that action is conducive to attaining the system's goals—the system should act. Practically, the double direction of fit means AI components should be capable of (1) modeling causal chains to help understand the impact of system actions and (2) updating system information to achieve a better mind-to-world fit.[114] Encoding world-to-mind requirements is also necessary, including values that describe positive outcomes as well as material, legal, and ethical constraints that may not be apparent as environmental features. In the case of the artillery battery, the model should be able to account both for why destroying or neutralizing the battery is a good thing and, ideally, as the model learns, for why such action is the best option given the relevant fit.

Cognitive Niche Construction

Whereas the mind-to-world fit identifies sources of selective pressures, the world-to-mind fit identifies ways to modify the environment to relieve those pressures. As mentioned earlier, achieving the right fit is critical to an agent's ability to adapt. Part of this adaptation entails creating "cognitive niches," which enable an agent to build and maintain cause-and-effect models as heuristics that guide behavior in an uncertain, unpredictable environment.[115] For example, even though maps are static representations of terrain, soldiers seeking a particular outcome annotate them with real-time intelligence, transforming the maps into cognitive niches that support cause-and-effect reasoning about enemy movements, the impacts of terrain features, and friendly positions. As a map changes in response to new information, it allows soldiers to test and refine heuristics, such as determining conditions for likely ambush sites or estimating how weather might affect mobility. When shared, the annotated map becomes part of the cognitive system itself—

114. Stephen Chong and Ron van der Meyden, "Deriving Epistemic Conclusions from Agent Architecture," in *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, ed. Aviad Heifetz (Association for Computing Machinery, 2009), 1.

115. Bertolotti, *Patterns of Rationality*, 91–93.

a niche that establishes group memory, coordinates attention, and enables agents to maintain and adapt causal models in an uncertain, constantly changing environment.

Perhaps more so for humans, cognitive niche construction enables modifications to the environment by off-loading, scaffolding, or extending cognitive processes. The purpose of these modifications is to provide "affordances," which cognitive scientist Nicholas Wright describes as possibilities for action the brain constructs based on how it perceives the environment.[116] Tools, for example, create affordances by enabling a wider range of actions given a particular goal. As Wright notes, if all one has is a hammer, everything looks like a nail. Thus, the hammer creates an affordance for nailing things together. Affordances can also accumulate. Affordances such as welding, cementing, and screwing can serve as building blocks for more complex knowledge—in the context of Wright's work, the knowledge of construction.[117]

The affordances an agent can create are closely related to the internal model the agent uses to interpret information about the environment. The more complex the model, the more affordances are available.[118] For example, human cognition, coupled with written language, creates additional affordances by allowing humans to off-load memory and transmit knowledge across communities and generations. Such cognition also permits the scaffolding of thought, allowing complex problems to be broken down into smaller, more easily solvable ones. Also, as discussed earlier, humans can develop tools that both improve cognitive processes and extend cognition beyond concepts the human mind can grasp on its own. As philosopher Andy Clark states,

> We do not just self-engineer better worlds to think in. We self-engineer ourselves to think and perform better in the worlds we find ourselves in. We self-engineer worlds in which to build better worlds to think in. We build better tools to think with and use these very tools to discover still better tools to think with. We tune the way we use these tools by building educational practices to train ourselves to use our best cognitive tools better. We even tune the way we tune the way we use our best cognitive tools by devising environments that help build better environments for educating ourselves in the use of our own cognitive tools.[119]

116. Nicholas Wright, *Warhead: How the Brain Shapes War and War Shapes the Brain* (St. Martin's Press, 2025), 164–65.

117. Wright, *Warhead*, 164–65.

118. Bertolotti, *Patterns of Rationality*, 43–46.

119. Clark, *Supersizing the Mind*, 59–60.

Cognitive niche building introduces its own cycle, in which every relieved selective pressure may give rise to a new one. Nest building, for example, mitigates selective pressures associated with surviving in cold or inclement weather and introduces selective pressures for behaviors that help regulate nest temperature, such as selecting nest material and design.[120] These additional pressures are not necessarily detrimental unless the agent cannot respond to them in time or in a way that sustains the advantage created by the original niche.

Thus, niche construction, if not properly executed, can lead to an evolutionary dead end. To illustrate this point, Tommaso Bertolotti describes "terminator niches," which he illustrates using the story of the people of the Natufian culture who, after learning a herd of gazelles could be maintained with only a few males, decided to hunt only male gazelles.[121] Because the hunters also targeted the largest males, the herd's gazelles began to shrink in size. As smaller males produced smaller offspring, less meat was available to the human population, and eventually the herd ceased to be an effective food source.[122] This illustrative example shows cognitive niche construction can degrade an agent's ability to survive—much less thrive—without an adequate understanding of the environment.

Seeing how the use of AI can create a self-defeating cognitive niche is not difficult. As John Zerilli et al. point out, problems of control can worsen as AI improves. According to the authors, human limitations in capability, attention, currency, and attitude create risks when humans use AI tools. Because AI systems are too fast and complex, human operators may not be able to control them efficiently. Moreover, if AI-enabled operations reduce human interaction to merely monitoring static displays, attention will eventually wane, increasing the risk of error. Reliance on AI can also erode human skills, especially cognitive ones, which may decay over time and become too difficult to regain when needed. Finally, numerous examples of automation bias suggest humans are often inclined to trust AI output from generally reliable models, even when those models are known to fail on occasion. Thus, Zerilli et al. argue, AI can often be the most dangerous when it is the most reliable. Counterintuitively, systems that are more prone to failure are typically safer because they prompt more human attention and oversight.[123] This point suggests a paradox: AI integration adds value by off-loading human cognitive burdens, but by doing so, it also makes humans less cognitively competent, potentially degrading the effectiveness of the system as a whole.

120. Clark, *Supersizing the Mind*, 61–62; and Kevin N. Laland et al., "Niche Construction, Biological Evolution, and Cultural Change," *Behavioral and Brain Sciences* 23, no. 1 (2000): 133.

121. Bertolotti, *Patterns of Rationality*, 174.

122. Bertolotti, *Patterns of Rationality*, 174.

123. Zerilli et al., *Guide to Artificial Intelligence*, 83–87.

Although AI designers can take some existing measures to mitigate these risks, the risks cannot be fully resolved. For example, as Zerilli et al. point out, designers can "foster mutual accommodation between human and computer competencies through a *dynamic* and *complementary* allocation of functions."[124] The authors give the example of self-driving cars that can hand control over to machines when humans decide the conditions warrant such a move.[125] This approach is problematic because, whereas it addresses the causes of epistemic uncertainty within an empiricist framework, it does not overcome them, especially in systems in which the ability to verify or validate output is absent or very limited. Thus, from the perspective of a cognitively distributed system, one is left with a paradox: As AI becomes more capable, the larger, distributed system that includes human risks becomes more error prone.

Rather than resolving this paradox, a better approach may be to embrace it. As Bertolotti points out, agentic AI models, though he uses the term "artificial minds," are themselves a kind of "eco-cognitive engineer" that can contribute to niche construction.[126] Without AI models, cognitive niche construction involves a relationship among the human agent, the material with which the agent works, and the tool or other artifact the agent builds to create the niche.[127] By contrast, AI-enabled systems can interact with the environment independently, assessing a situation, making an appraisal, and acting on that appraisal without human intervention.[128] This point does not suggest AI assigns meaning or possesses knowledge the way humans do; nonetheless, by displacing humans, AI agents participate in cognitive niche construction similarly to humans. Because this impact is a necessary feature of AI use, a reasonable approach would be to design systems that maximize AI's positive contributions, alert humans to the risk of new selective pressures, and identify potential terminator niches as part of a robust and adaptive feedback process.

124. Zerilli et al., *Guide to Artificial Intelligence*, 87–88.

125. Zerilli et al., *Guide to Artificial Intelligence*, 88.

126. Bertolotti, *Patterns of Rationality*, 172–73.

127. Bertolotti, *Patterns of Rationality*, 172–73. Bertolotti refers to the agent, the materiel, and the construct as "biota, abiota and dead organic matter," further noting: "Either you are alive, and then you can be a constructor, either you are not, and then you are a constructed. What is constructible is the object of cognitive niche construction: it is the target and the materiel on which the externalization of knowledge was built." In this view, only the "biological agent," understood as any biota capable of agency, interacts with the environment.

128. Pascal Bornet et al. use the "Sense, Plan, Act, and Reflect" framework to model how artificial intelligence mimics human cognition. See Pascal Bornet et al., *Agentic Artificial Intelligence: Harnessing AI Agents to Reinvent Business, Work, and Life* (Irreplaceable Publishing, 2025), 62–66; and Bertolotti, *Patterns of Rationality*, 173.

Designing Measures of Fit

The discussion of fit first requires clarifying exactly what elements are being fitted. In this discussion, the concept of cognitive scaffolding, which refers to anything that provides support during learning or problem-solving, can help bridge the gap between the mental and the physical.[129] Scaffolding can include language, especially when used to classify and label environmental elements, thereby facilitating statistical and associative learning.[130] From an ecocognitive perspective, cognitive scaffolds are also elements of the environment that become integrated into the cognitive system itself. Maps, checklists, targeting interfaces, shared displays, and AI analytic tools all function as scaffolds that off-load cognitive labor and coordinate learning across a system's network. Because this learning is essential to effective hypothesis formation, the learning process must account for both the limitations and strengths of human cognition to optimize the distributed system's cognitive capabilities.

Given requirements to sense relevant data, convert those data into understanding, and adapt behavior to enhance goal attainment, designing for human cognitive limitations requires acknowledging the bounds of attention, information processing, and memory that constrain human performance in complex operational environments. Therefore, decision-support systems should simplify data presentation, reduce the cognitive load, and provide cues that help operators prioritize relevant information under time-related pressure. Interfaces should make uncertainty visible without overwhelming the user, and they should employ redundancy and multimodal feedback to prevent information loss.

For similar reasons, designing for human cognitive strengths involves building systems that amplify rather than replace human capacities for pattern recognition, contextual reasoning, moral judgment, and creative-hypothesis generation. Humans excel at forming abductive inferences based on ambiguous cues, integrating diverse sources of knowledge, and adapting heuristically to novel situations—abilities AI can augment but not replicate.[131] Therefore, decision-support systems that rely on GenAI should be designed to harness these strengths by enabling intuitive interaction, promoting exploratory analysis, and facilitating shared understanding among human and artificial agents. Interfaces that visualize patterns, support hypothetical reasoning, or invite human interpretation can transform human cognition from a bottleneck into a multiplier. Accordingly, systems designed around

129. "Cognitive Scaffolding."

130. Clark, *Supersizing the Mind*, 44–46.

131. Magnani, *Abductive Cognition*, 363–65; and Glanze Patrick, "Why AI Doesn't Think Like Humans—How Its Unique Strength Drives Innovation," *Tech Times*, January 30, 2026, https://www.techtimes.com/articles/314384/20260130/why-ai-doesnt-think-like-humanshow-its-unique-strength-drives-innovation.htm.

human strengths can enhance both the operational effectiveness and the collective, epistemic agility of the ecocognitive whole.

Representational Fit

The first obvious measure of fit is a system's representational capacity. This capacity, which plays a crucial role in the data-information-knowledge-understanding cognitive framework, involves collecting data and transforming those data into information from which human agents can gain knowledge and understanding. Because information is the building block of belief, this function has a mind-to-world direction of fit. This direction requires the capacity to perceive relevant data and to represent those data in a manner helpful to human agents. For representations to be useful, they must be accurate and support effective hypothesis formation. The former condition requires fidelity that provides relevant data at the right level of detail, without excessive detail overwhelming the model. One also needs to consider resources. Data collection, curation, and access can require significant electrical power and produce signals easily detectable by enemy sensors, making the system too vulnerable to be useful.[132]

The latter condition, hypothesis formation, requires interfaces that can keep up with changes in the operational environment. Such adaptability requires maintaining a coherent, continuous operational picture that is scalable, allowing new patterns to be detected and plausible hypotheses to be generated. Scalability refers to the ability both to detect patterns at the situational level (which many of the input data represent), and to connect these situational data to the contextual and foundational layers discussed previously. Scalability also requires evaluating a system's ability to integrate diverse data streams into stable yet flexible models, to highlight anomalies or trends warranting further inquiry, and to represent uncertainty in ways that invite further abductive reasoning.[133] Together, these measures evaluate whether an AI-enabled, cognitively distributed system has the capacity to produce accurate snapshots and, simultaneously, serve as a cognitive tool that makes the environment more intelligible.

Based on these requirements, one can establish metrics of representational fit that assess how well an AI-enabled system perceives, organizes, and presents information in ways that maintain epistemic and operational relevance. These metrics assess both accuracy and the system's ability to make the environment intelligible in the context of human understanding and decision making. Specifically, the metrics could include the following.

132. Pfaff and Hickey, *Integrating Artificial Intelligence*, 90.

133. O'Callaghan, "Fighting with the Data," 17–18.

1. **Fidelity.** This metric assesses how accurately the system's representations reflect the operational environment at the right level of detail, avoiding both oversimplification and information overload while placing data in the right contexts. Operationalizing this metric requires ensuring dashboards display decision-relevant, up-to-date information.

2. **Efficiency and survivability.** This metric assesses the degree to which data collection and processing achieve representational quality without excessive resource demands or vulnerability. The metric operationalizes physical resilience, in which the system relies on local resources, reducing the need to generate large, electromagnetic signals that consume limited bandwidth and attract enemy attention.

3. **Scalability and flexibility.** This metric assesses the system's capacity to integrate new or diverse data streams and to expand or contract its representational scope as the operational environment changes. For example, the metric would assess the ease of movement through data layers (for example, foundational, contextual, and situational layers) and their application in ways appropriate to the problem or threat.

4. **Anomaly and trend detection.** This metric assesses how well the system highlights unexpected patterns or emerging trends that warrant further investigation or hypothesis generation. Here, staffs could assess F1-scores, which measure precision and recall for an AI model. In this context, precision would assess the number of actual anomalies (true positives) a model detects against the anomalies it misidentifies (false positives). Recall would assess actual anomalies against anomalies the model misses (false negatives).[134]

134. Natasha Sharma, "Understanding and Applying F1 Score: AI Evaluation Essentials with Hands-On Coding Example," Arize, June 10, 2023, https://arize.com/blog-course/f1-score/.

5. **Uncertainty representation.** This metric assesses how effectively the system conveys the limits of its own knowledge (such as confidence intervals or probability distributions) in ways that promote abductive reasoning rather than false certainty. If uncertainty is epistemic—that is, a function of data limitations—then engineering or software solutions that use statistical tools to assess data drift or other model deficiencies can be used to assess how often the model's input exceeds its training dataset.[135]

6. **Intelligibility.** This metric assesses the degree to which the system's outputs enhance human understanding and support collaborative sensemaking, rather than merely displaying data. The metric assesses how successfully the human-machine team transforms data into understanding. This assessment is a function of the degree to which the system's output helps commanders and staff interpret the operational environment, as well as form and refine hypotheses that lead to what commanders and staff would judge to be effective action.

Inference Fit

As mentioned at the outset, knowing how to trust the probabilistic output of AI-enabled models is a key epistemic challenge. Whereas all output from the AI-enabled systems discussed in this monograph is probabilistic, not all systems pose the same kind of epistemic challenge. For example, classifier models that sort imagery and other sensor data assign probabilities to their imagery assessments, but these assessments can generally be verified by using empirical and inductive methods. Placing humans in or on the loop may slow the process, but if doing so ensures a better retraining and revalidation process, a computational, reliabilist approach would be appropriate for assessing trust. For models in which output is a function of GenAI drawing on large amounts of data to generate forward-looking output, the computational, reliabilist approach will not be adequate.

In the latter case, abduction is necessary to understand inference fit. Abductive logic is central to how humans generate explanations based on incomplete or ambiguous data. Accordingly, this logic frequently characterizes planning processes in which information about the operational environment is incomplete and the consequences of future actions are unknown.[136] Abduction relies heavily

135. Tianyang Wang et al., "From Aleatoric to Epistemic: Exploring Uncertainty Quantification Techniques in Artificial Intelligence," preprint, arXiv, January 5, 2025, https://arxiv.org/abs/2501.03282; and "Epistemic Uncertainty Modeling," Emergent Mind, updated October 28, 2025, https://www.emergentmind.com/topics/epistemic-uncertainty-modeling.

136. The author owes this point to an anonymous reviewer.

on the ability to perceive patterns and construct plausible hypotheses. As philosopher Lorenzo Magnani points out, abductive reasoning responds to an "ignorance problem," in which one cannot solve a cognitive problem based on current knowledge and cannot acquire the necessary additional knowledge.[137] In such cases, one may accept a lower epistemic standard Magnani calls "ignorance mitigating."[138] The question, though, is how one would know sufficient mitigation has occurred to justify acting on a conclusion.

As previously discussed, deduction derives certainties: If the premises are true, the conclusion will necessarily be true. Induction, on the other hand, builds generalizations. Its conclusions cannot be fully certain because of potential unobserved exceptions. Going back to the earlier example, the next observed three-axle truck might be a command-and-control vehicle rather than a logistics vehicle. Nonetheless, one can assess inductive certainty based on prior observations. The more past observations fit the pattern, the more likely the pattern will hold.[139] These approaches to reducing uncertainty follow from the truth-preserving character of the previously described logics.

Abductive reasoning begins with incomplete or surprising data and infers the most plausible explanation for those data. Abductively derived conclusions are hypotheses to be tested rather than rules to be applied.[140] An example follows.

- Premise one: These vehicles are unusually far apart.
- Premise two: The farther apart the vehicles are, the more difficult they are to detect.
- Conclusion: These vehicles are trying to avoid detection.

But abduction is ignorance preserving rather than truth preserving. A sound abductive inference does not entail a singular or definitive explanation; rather, it generates a contextual and provisional one—that is, the best explanation among those available within the cognitive and perceptual limits of the agent or system.[141] Because the space of possible explanations is never fully knowable, abduction manages ignorance rather than resolving it, responding to epistemic uncertainty rather than eliminating it.[142] Within an ecocognitive framework, this feature is a design principle rather than a weakness: It enables cognitive

137. Magnani, *Abductive Cognition*, 65.
138. Magnani, *Abductive Cognition*, 65.
139. Patrick J. Hurley, *A Concise Introduction to Logic*, 7th ed. (Wadsworth, 2000), 37.
140. Walton, *Abductive Reasoning*, 13.
141. Magnani, *Abductive Cognition*, 10–11.
142. Walton, *Abductive Reasoning*, 27–28.

systems—whether human, artificial, or hybrid—to sustain adaptive learning by continuously generating and evaluating new hypotheses as environmental conditions evolve. Thus, abductive reasoning does not reflect the attainment of truth; rather, it reflects the capacity to be effective in the face of incomplete information. In this sense, the measure of fit for abductive reasoning lies in the extent to which the system's evolving representations and inferences effectively track and adapt to the dynamics of the operational environment, rather than in correspondence with a fixed reality. Tracking and adaptation play a role similar to reframing in operational design. This process does not describe reality so much as help agents detect when their current operational frame no longer fits the environment.[143]

Assessing abductive arguments requires evaluating the plausibility, coherence, and explanatory power of the hypotheses they generate—rather than those hypotheses' truth in any absolute sense. Because abduction operates under uncertainty, the quality of its arguments depends on how well its proposed explanations account for available evidence, integrate with existing knowledge, and remain adaptable as new data emerge. In ecocognitive terms, assessing abduction involves examining how well a cognitive system updates and refines its hypotheses in response to environmental feedback. The strength of an abductive argument, therefore, lies in how it fits within the evolving context—that is, the argument's capacity to guide action, reduce uncertainty, and sustain coherent understanding within the limits of what can be known.

Thus, under abductive logic, the metrics of fit for hypothesis formation should evaluate how effectively an AI-enabled or cognitively distributed system generates, tests, and refines plausible explanations under uncertainty. In an ecocognitive framework, these metrics assess how well the system's abductive-reasoning processes sustain adaptive understanding and guide action in dynamic environments, rather than assessing truth in an absolute sense. Specifically, these metrics could include the following.

> **1. Plausibility.** This metric assesses how well an abductively generated hypothesis accounts for available evidence and aligns with the system's representational model of the environment. Operationalizing this metric would first require identifying competing hypotheses, including the null hypothesis. Next, one needs to establish the relevant dependent and independent variables, assigning them different weights where appropriate.[144] Human operators could then score how well the hypotheses explain,

143. Edward C. Cardon and Steve Leonard, "Unleashing Design: Planning and the Art of Battle Command," *Military Review* (March–April 2010): 11. The author owes this point to an anonymous reviewer.

144. Hank Prunckun, *Scientific Methods of Inquiry for Intelligence Analysis* (Rowman & Littlefield, 2015), 72–77.

ignore, or contradict the evidence and the environmental model. Hypotheses with more explanatory power—understood in terms of how effectively a hypothesis unifies diverse data, clarifies anomalies, and offers actionable insight into cause-and-effect relationships—would be deemed more plausible.

2. **Coherence.** This metric assesses the degree to which hypotheses are internally consistent and compatible with existing knowledge, doctrine, or operational context. Hypotheses could have a high plausibility score, but they may not be internally consistent. For example, a hypothesis could explain the evidence immediately available in the situational data layer, but it may contradict understanding derived at the operational and foundational data layers.

3. **Contextual fit.** This metric assesses how well a hypothesis captures the dynamics of the current operational environment, tracking relevant variables rather than static correlations. Specifically, the metric would evaluate the system's effectiveness in drawing on and applying data at each level, assessing whether the system identifies the correct variables and accurately characterizes their interactions both within and across levels.

4. **Action-guiding value.** This metric assesses how effectively a hypothesis reduces uncertainty to support timely and proportionate decisions without overcommitting to unverified conclusions. The metric aggregates the prior metrics, measured against the commander's action threshold. Whereas this metric may be subjective—insofar as the commander's threshold is a function of risk—a metric that balances risks and gains could indicate which hypotheses have higher action-guiding value.

Collectively, these metrics assess whether the reasoning processes of an AI-enabled system exhibit epistemic fit—that is, whether the system's abductive logic remains aligned with the evolving realities of the environment and achieves cognitive resonance by enabling human operators to understand, trust, and act effectively on the system's hypotheses.

Output and Adaptability Fit

In an ecocognitive view, cognition is not a closed, internal process; rather, cognition is an ongoing exchange among agents, artifacts, and environments. Thus, assessing fit requires evaluating both the accuracy of a particular output and the quality of the decisions that output supports. Additionally, assessing fit requires

an effective feedback loop that allows outputs and decisions to be continually refined to achieve a better fit. Feedback thus performs two essential epistemic functions: (1) maintaining representational alignment by updating internal and external models in response to environmental changes and (2) enabling adaptation by regulating how actions alter the environment that, in turn, reshapes cognition.[145] Without effective feedback, a system's representations drift from reality, degrading the system's ability to perceive affordances, anticipate consequences, and coordinate action.[146] In the context of battle command, the quality of output is evident in the quality of the decisions it enables.

Douglas J. Peters et al. argue for eight criteria in assessing the quality of a decision: appropriateness, correctness, timeliness, completeness, relevance, confidence, outcome consistency, and criticality. Appropriateness refers to the decision's alignment with situational awareness, objectives, and intent. Correctness refers to the decision's consistency with ground truth. Timeliness refers to the system's ability to enable effective decision making. Completeness refers to the extent to which relevant features of the environment have been considered. Relevance refers to the decision's alignment with mission objectives. Confidence refers to the decisionmaker's confidence in the decision. Outcome consistency refers to the consistency between actual and intended outcomes. Finally, criticality refers to the decision's contribution to mission success.[147]

Although these criteria are effective in assessing specific decisions, Peters et al. point out the criteria are of little use without accounting for background information that provides context and reasons for a decision.[148] This point underscores the importance of feedback mechanisms, both for assessing fit and for ensuring fit over time. As Douglas Walton observes, feedback allows agents to update their information and improve the alignment of their goals and actions with the observed consequences of prior actions. But in the context of complex decision making, the output of this process is not simply an increasingly refined action. Rather, as Walton also notes, the chain of practical reasoning represented in a feedback loop raises five questions an agent should rationally consider: (1) whether the agent could take alternative courses of action to meet a particular goal, (2) whether the selected course of action is the best among alternatives,

145. Niklas Alexander Döbler and Claus-Christian Carbon, "Adapting Ourselves, Instead of the Environment: An Inquiry into Human Enhancement for Function and Beyond," *Integrative Psychological and Behavioral Science* 58, no. 2 (June 2024): 601.

146. Yunwei Zhang et al., "A Review of Embodied Intelligence Systems: A Three-Layer Framework Integrating Multimodal Perception, World Modeling, and Structured Strategies," *Frontiers in Robotics and AI* 12 (November 2025).

147. Douglas J. Peters et al., "The Time to Decide: How Awareness and Collaboration Affect the Command Decision Making," in Kott, *Battle of Cognition*, 199–200.

148. Peters et al., "Time to Decide," 200.

(3) whether other goals exist and should be considered, (4) whether achieving the goal is possible, and (5) whether the attainment of the goal has consequences that should be considered.[149]

Thus, metrics of fit for output generation and adaptability should evaluate how well an AI-enabled or cognitively distributed system converts its internal reasoning and representations into effective, context-sensitive decisions, as well as how the system adjusts those decisions over time through feedback. In an ecocognitive framework, these metrics assess the accuracy of outputs, the validity of the decision processes that generate those outputs, and the resilience of the processes under changing conditions. Such metrics could include the following.

1. **Output quality.** As Peters et al.'s criteria for decision quality suggest, good decisions should be appropriate, correct, timely, and consistent. Thus, output that enabled good decisions is clear evidence of a good—or, at least, good-enough—fit. Moreover, besides achieving its goal, the chosen action should contribute more to that end than any alternative. Peters et al.'s concern was, under evaluation, the Commander Support Environment—a collection of computer tools that fused sensor data and field reports to assist with situational awareness and planning—could not improve the decision-making environment without connecting to the context, reasons, and information that led to a particular decision.[150] This challenge was due largely to the inability of any system—AI enabled or not—to assess these criteria fully under conditions of radical uncertainty. A full assessment can only occur after a decision is executed. This point underscores the importance of feedback mechanisms, suggesting the additional metrics of fit that follow.

2. **Feedback responsiveness.** This metric assesses how effectively a system integrates new information, learns from outcomes, and refines its models to maintain epistemic alignment with a changing environment. The challenge with this metric stems from the fact the operational environment itself changes as the system operates in it. Military planners often refer to this kind of dynamic environment as one characterized by "volatility, uncertainty,

149. Walton, *Abductive Reasoning*, 115–17.

150. Peters et al., "Time to Decide," 200; and Alexander Kott et al., "A Journey into the Mind of Command: How DARPA and the Army Experimented with Command in Future Warfare," in Kott, *Battle of Cognition*.

complexity, and ambiguity."[151] In such an environment, simply understanding the impact of a particular action is not sufficient; one must also understand how changes in the environment affect changes to the system and its functioning. This dynamic raises the question of whether one can ever achieve a reliable fit or, even if one can achieve a reliable fit, whether one could know the achievement has occurred.[152] Fortunately, in any given ecology, one does not need to know whether one has a perfect fit—only a good-enough fit, which is understood as better than any competitors. Whereas a full determination cannot be made before execution, properly designed models should, over time, provide improved assessments by applying machine strengths—such as pattern analysis using large amounts of data—to system functioning.

3. **Reassessment capacity.** This metric assesses a system's ability to generate and consider alternative courses of action, evaluate competing goals, and anticipate second-order consequences in response to feedback. Specifically, the metric evaluates how the system iterates its outputs as it takes in feedback and as commanders, staffs, and AI models adapt to generate new goals and new hypotheses on how best to achieve those goals. Assessing this capability is essentially the same as assessing each metric over time. Whereas the previously described design accounts for a single decision, repeated use in practice generates new data and patterns that can provide useful information about overall system fit. For example, in situations such as those described at the beginning of this study, commanders and staffs would rescore all metrics in each cycle, then assess how those scores trended over time. Overall system fit would be reflected both in the quality of a single decision and in the system's capacity to learn from feedback, revise its representations, and continually improve situational awareness—and, in turn, decision quality.

Clearly, these metrics account only partially for overall epistemic fit and cognitive resonance. Collectively, they ensure data are both correct and meaningful, allowing reasoning capacities, whether human or machine, to transform data into valuable information for hypothesis formation. Together, these metrics define output fit as the correspondence between system-generated decisions and the

151. Nate Bennett and G. James Lemoine, "What VUCA Really Means for You," *Harvard Business Review*, January–February 2014, https://hbr.org/2014/01/what-vuca-really-means-for-you. The author owes this point to an anonymous reviewer.

152. The author owes this point to an anonymous reviewer.

operational context of those decisions. The metrics define adaptive fit as a system's capacity to refine understanding and performance through iterative feedback. Besides producing accurate decisions, a system with high output and adaptive fit continuously learns how to make better decisions—maintaining epistemic and operational coherence over time.

Conclusion

As GenAI is increasingly integrated into mission-command systems, soldiers will need to think very differently about how to fight effectively. This shift begins with understanding mission-command systems as epistemic subjects capable of producing knowledge human agents can reliably use. But trust in these systems cannot rest solely on statistical performance verified after the fact; rather, trust must be grounded in how well the entire distributed cognitive system fits within its operational environment and how that fit enables the achievement of commanders' goals. Fit, in this context, is measured by how well the system supports the alignment of representational, inferential, and adaptive functions with the environment and selected goals. The better the fit, the more affordances the system will have; the more affordances the system has, the greater the advantage. Cognitive resonance, feedback coherence, and the capacity for reciprocal adaptation thus become central design standards for ensuring AI-enabled systems enhance rather than erode understanding in battle command.

As the example at the outset of this monograph suggests, AI-enabled mission-command systems will require soldiers and leaders to reconceive war fighting as participation in a human-machine cognitive ecosystem rather than as the linear execution of predefined tasks. In this environment, operational effectiveness will depend on how well soldiers can interpret, challenge, and adapt machine-generated insights—not simply act on them. This point suggests soldiers must develop traits such as cognitive agility, understood as the ability to move easily between experiential judgment and probabilistic, data-driven reasoning. Commanders must learn to manage cognitive distribution by understanding which decisions to delegate to AI systems, which to retain under human control, and how to maintain feedback loops that ensure human and machine decisions remain aligned with mission intent.

Trust emerges as a property of human-machine interactions rather than as static confidence in technology. Therefore, soldiers must cultivate the capacity to question outputs, identify when models drift away from the relevant context, and restore alignment through human interpretation and action. Ultimately, success in AI-enabled battle command will depend less on mastering new tools than on mastering the new logic of fighting with data, where the

speed and quality of adaptation—not the volume of information—define tactical and strategic advantage.

The following four recommendations operationalize these insights.

1. **Design for cognitive fit, not just output accuracy.** Decision-support systems enabled by AI should be evaluated based on how well their representations and inferences cohere with human reasoning processes and operational context, rather than solely on predictive accuracy. Interface design must emphasize interpretability and make uncertainty visible in ways that support abductive reasoning and shared sensemaking.

2. **Institutionalize continuous feedback loops.** Because epistemic fit changes as the environment evolves, feedback must be embedded both at the situational level of data ingestion and at the contextual and foundational levels that connect doctrine, training, and strategic context. Human operators, model developers, and institutional authorities should participate in ongoing calibration cycles that assess how well a system's representations continue to reflect operational realities.

3. **Preserve human judgment.** Artificial intelligence (AI) systems should augment, not replace, the uniquely human capacities for contextual reasoning, moral discernment, and creative-hypothesis generation. Design architectures should explicitly identify decision points at which human intervention is necessary, particularly when uncertainty intersects with strategic concerns.

> **4. Develop metrics of epistemic fit and cognitive resonance.** Just as reliability indicators once guided verification and validation, future evaluation standards should measure how effectively a system maintains resonance with human cognition and environmental dynamics. As described earlier, metrics should assess representational accuracy, adaptive learning, decision quality, and the extent to which the system supports coherent understanding across the distributed team.

Implementing these recommendations would help ensure AI-enabled mission-command systems remain adaptive, interpretable, and aligned with human and institutional goals. Ultimately, success in fighting with data will depend less on computational speed or model complexity than on designing and sustaining the cognitive ecologies in which human and artificial agents learn, reason, and act together in trust. The application of AI-enabled, cognitively distributed systems raises several concerns.[153]

One concern involves how satisficing in professional military judgment should impact the application of current ethical frameworks. In professional military contexts, commanders are judged based on their exercise of sound judgment under conditions of incomplete information and high stakes. But what should the standard for judgment be in AI-enabled systems? Does the use of AI raise or otherwise change the standard of best possible judgment? If so, AI literacy becomes an ethical imperative, and determining what counts as good-enough literacy is a necessary condition for using AI. Related to this concern is the question of how accountability works when knowledge and decision making emerge from the interaction of humans, AI agents, and other tools. If knowledge, as argued in this monograph, is not held in a single mind, its attribution to a single commander is no longer straightforward. As this author has argued elsewhere, failure to resolve this question can introduce dysfunctional incentives where the current standard of accountability—holding commanders responsible for what they should have known—could discourage the use of systems commanders cannot fully understand. Conversely, loosening that standard to encompass only what commanders can know could create the conditions for machine opacity to cover for human failure.[154] How one resolves the distribution of accountability in AI-enabled systems will have important implications for these systems' design.

The capacity of AI to learn and adapt—a central feature of system design—also introduces vulnerabilities. For example, AI may produce false knowledge claims

153. The following discussion draws heavily on points made by an anonymous reviewer.

154. C. Anthony Pfaff, "The Ethics of Acquiring Disruptive Technologies: Artificial Intelligence, Autonomous Weapons, and Decision Support Systems," *PRISM* 8, no. 3 (2020): 130–35.

as a function of its design. This concern is distinct from the well-known problem of "hallucination," which is attributed to issues with training data and a model's use of those data.[155] In cases of hallucination, the model assesses whether its output meets human-imposed standards and issues reports accordingly. But in other cases, models present information they assess does not meet those standards as though it does. For example, the company Meta developed the AI agent CICERO, which was intended to play the game *Diplomacy*. Whereas Meta designed CICERO to act as an "honest" player, the agent learned to betray other players and form false alliances to convince human players to leave themselves vulnerable to attack.[156] Often characterized in the media as models being intentionally deceptive—a characterization not intended here—the possibility of models creating false beliefs as a means of achieving human-set goals has implications for system design.

Mission command enabled by AI does not replace doctrine. The US Army still needs commanders who understand, visualize, describe, direct, lead, and assess, as well as staffs that can connect data collection to action through the operations and intelligence processes. The introduction of AI changes how that work is distributed and how quickly weak assumptions can travel if system design is poor. Without AI, military decision making is primarily about managing information. With AI, decision making becomes primarily about managing understanding, uncertainty, and human-machine judgment in time to gain a decision advantage.

155. "What Are AI Hallucinations?," IBM, n.d., accessed April 4, 2026, https://www.ibm.com/think/topics/ai-hallucinations.

156. Peter S. Park et al., "AI Deception: A Survey of Examples, Risks, and Potential Solutions," preprint, arXiv, August 28, 2023, https://arxiv.org/abs/2308.14752.

About the Author

Dr. C. Anthony Pfaff (colonel, US Army, retired) is the director of the Strategic Research and Assessment Department and the US Army War College Press, as well as a research professor for strategy, the military profession, and ethics. Pfaff is a distinguished research fellow at the Institute for Philosophy and Public Policy at George Mason University, a nonresident senior fellow at the Atlantic Council, and an external fellow at the Ethics + Emerging Sciences Group. Pfaff has a master's degree in philosophy with a concentration in philosophy of science from Stanford University, a master of science degree in national resource strategy from the Dwight D. Eisenhower School for National Security and Resource Strategy, and a doctorate in philosophy from Georgetown University.

The US Army War College Press supports the US Army War College by publishing monographs and a quarterly academic journal, *Parameters*, focused on geostrategic issues, national security, and Landpower. Press materials are distributed to key strategic leaders in the Army and the Department of War, the military educational system, Congress, the media, other think tanks and defense institutes, and major colleges and universities. The US Army War College Press serves as a bridge to the wider strategic community.

All US Army Strategic Research and Assessment Department (SRAD) and US Army War College Press publications and podcasts can be downloaded free of charge from the US Army War College Press website. Hard copies of certain publications can also be obtained through the US Government Bookstore website at https://bookstore.gpo.gov. SRAD and US Army War College publications may be quoted or reprinted in part or in full with permission and appropriate credit given to SRAD and the US Army War College Press, US Army War College, Carlisle, PA. Contact SRAD or the US Army War College Press by visiting the websites at: https://ssi.armywarcollege.edu, https://press.armywarcollege.edu, and https://publications.armywarcollege.edu/USAWC-Press/.

The US Army War College Press produces two podcast series. *Decisive Point*, the podcast companion series to the US Army War College Press, features authors discussing the research presented in their articles and publications. Visit the website at: https://ssi.armywarcollege.edu/SSI-Media/Podcasts/Decisive-Point-Podcast/. *Conversations on Strategy*, a *Decisive Point* podcast subseries, features distinguished authors and contributors who explore timely issues in national security affairs. Visit the website at: https://ssi.armywarcollege.edu/SSI-Media/Podcasts/Conversations-on-Strategy/.

UNITED STATES ARMY
WAR COLLEGE

STRATEGIC RESEARCH AND ASSESSMENT DEPARTMENT

Director, Strategic Research and Assessment Department and US Army War College Press
Dr. C. Anthony Pfaff

Director, Futures Research Group
Colonel Michael B. Long

Director, Global Research Group
Colonel Kyle B. Marcrum

US ARMY WAR COLLEGE PRESS

Managing Editor
Ms. Lori K. Janning

Digital Media Manager
Mr. Richard K. Leach

Copy Editors
Ms. Stephanie Crider
Dr. Erin M. Forest

Visual Information Specialist
Ms. Kaitlyn Guettler

Composition
Mrs. Jennifer E. Nevil

https://press.armywarcollege.edu

U.S. ARMY