Strategy Research Project

Integrating Artificial Intelligence into Military Operations: A Boyd Cycle Framework

by

Dr. James Mancillas
Department of the Army

Under the Direction of:
Dr. Gregory L. Cantwell

United States Army War College
Class of 2017

DISTRIBUTION STATEMENT: A
Approved for Public Release
Distribution is Unlimited

The views expressed herein are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Army, Department of Defense, or the U.S. Government. The U.S. Army War College is accredited by the Commission on Higher Education of the Middle States Association of Colleges and Schools, an institutional accrediting agency recognized by the U.S. Secretary of Education and the Council for Higher Education Accreditation.

REPORT DOCUMENTATION PAGE	*Form Approved--OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 01-04-2017	**2. REPORT TYPE** STRATEGY RESEARCH PROJECT	**3. DATES COVERED** *(From - To)*

4. TITLE AND SUBTITLE Integrating Artificial Intelligence into Military Operations: A Boyd Cycle Framework	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Dr. James Mancillas Department of the Army	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Dr. Gregory L. Cantwell	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army War College, 122 Forbes Avenue, Carlisle, PA 17013	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT Distribution A: Approved for Public Release. Distribution is Unlimited.
To the best of my knowledge this SRP accurately depicts USG and/or DoD policy & contains no classified information or aggregation of information that poses an operations security risk. **Author:** ☒ **PA:** ☒

13. SUPPLEMENTARY NOTES
Word Count: 7243

14. ABSTRACT

The emergence of artificial intelligence (AI) offers opportunities to overcome cognitive limits when managing “big data” and further leverage informational superiority into battle space superiority. However, these advantages may only materialize if these future technologies are properly integrated into military operations. To understand and facilitate this integration, this paper establishes baseline definitions and descriptions for key terms such as AI, autonomy, and degrees of autonomy, in-the-loop, and decision loops. Applying these terms, this study establishes a framework based on the Boyd cycle, also known as the Observe, Orient, Decide, and Act loop (OODA Loop), to explore how AI systems participate in military decision processes. For each part of the OODA Loop, this study briefly explores the implications of AI systems using the JCIDS DOTMLPF lens.

15. SUBJECT TERMS
OODA Loop, Autonomy

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES** 33	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT UU	**b. ABSTRACT** UU	**c. THIS PAGE** UU	UU		**19b. TELEPHONE NUMBER** *(w/ area code)*

Integrating Artificial Intelligence into Military Operations: A Boyd Cycle Framework

(7243 words)

Abstract

The emergence of artificial intelligence (AI) offers opportunities to overcome cognitive limits when managing "big data" and further leverage informational superiority into battle space superiority. However, these advantages may only materialize if these future technologies are properly integrated into military operations. To understand and facilitate this integration, this paper establishes baseline definitions and descriptions for key terms such as AI, autonomy, and degrees of autonomy, in-the-loop, and decision loops. Applying these terms, this study establishes a framework based on the Boyd cycle, also known as the Observe, Orient, Decide, and Act loop (OODA Loop), to explore how AI systems participate in military decision processes. For each part of the OODA Loop, this study briefly explores the implications of AI systems using the JCIDS DOTMLPF lens.

Integrating Artificial Intelligence into Military Operations: A Boyd Cycle Framework

This study examines the implications of artificial intelligence on future military operations. This study first presents definitions of key terms, to form a common baseline for later discussions. Then, it explores the integration and implications of artificial intelligence by mapping artificial intelligence functions onto military decision processes. These discussions are driven by: (1) the estimation of erosion of U.S Military overmatch with peer competitors; (2) the rapid emergence of revolutionary technologies, here to previously only existing in the realm of science fiction; (3) the rapid urbanization of societies around the globe and subsequent impact on future combat environments and the character of war; and, (4) differing social and political perspectives bounding the ethical limits for their implementation.

The U.S military has been a leader in implementing an array of emerging and revolutionary information age technologies. The ever-growing use of autonomous vehicles is an obvious example of these emerging technologies. The use of Unmanned Aerial Vehicles (UAV) or Remotely Piloted Vehicles (RPA) as Intelligence, Surveillance and Reconnaissance (ISR) platforms has given U.S forces unprecedented situational awareness. This enhanced understanding of the battle space allows Commanders to optimize forces and reduce risks by placing forces at the right location, at the right time, and in the right size.

However, there are early indicators that the adoption of these maturing information age technologies have yet to reach their full potential. In support of this conclusion, two obvious issues arise. The first concerns the cost and manpower required to support today's UAVs, and the second involves information saturation at the command and staff level.

The use of UAVs has improved military capabilities while reducing risks to personnel. However, these improvements have come at a cost. Consistent with the preceding years, President Obama's FY17 defense budget request included 4.61 billion dollars for drone-related spending. Additionally, at a time when there has been increasing pressure to reduce military manning, the U.S. Air Force estimates: that a MQ-1 Predator requires a crew of 168; a MQ-9 Reaper requires 180; and, a RQ-4 Global Hawk requires 300.[1] Furthermore, these "unmanned" platforms require system operators with expertise that has been extremely difficult to retain.[2]

The second issue indicating a misalignment of capabilities is the disparity between the ability to gather information and the ability to process that information into useful situational awareness. Since the demonstrated success of Network Centric Warfare (NCW), warfare with a focus on information superiority, there has been an ongoing quest to gather more information to gain or maintain "self-synchronized" battle space advantage.[3] This focus on information superiority has manifested itself in today's fleet of UAVs and integrated Command and Control (C2) systems. Yet in an operational as well as technological environment, the ability to contextualize this information has not keep pace with the technologies used to gather, store, and share information.

Processing and contextualizing information are more difficult tasks than gathering data. Metcalfe's Law describes the difficulty of this challenge.[4] In short, Metcalfe's Law ascertains that contextualizing information increases in difficulty as more data nodes are considered. The proportionality of this relationship is as a square of the number of data nodes (i.e., difficulty in deriving value from information = [number of information sources]^2).[5] For example, if the number of data sources increased by a factor of 3, the difficultly in contextualizing the information increases by a factor of 9. As a result, it is increasingly difficult to derive additional value from more information. However, computer technologies are starting to emerge that have the potential to dramatically change the calculus on information usability, and thereby considerably alter the character of war in the information age.

Until now, the ability to gather, store, and communicate information has been the defining feature of the information age. However, as the information age matures, the ability to process and distil information may be its new defining feature. Moreover, the ability to fully integrate information collection, communication, storage, and processing into timely and decisive action may result in technologically based military offsets. The technologies embodied in artificial intelligence (AI) and the development of autonomous systems may cumulatively result in what may be called the “third offset”.

However, achieving a third offset, a decisive military advantage based on the application of a new technology, is not a foregone conclusion. This paper will explore the implementation of these emerging technologies by first establishing some common definitions. Then, it establishes a unifying framework that integrates these concepts into a military decision process. And finally, using the Joint Capabilities Integration Development System (JCIDS) framework, this paper explores potential pitfalls and challenges to consider as the Department of Defense begins integrating these new technologies into military operations.[6]

Definitions

Discussions about the rapidly evolving landscape of autonomous systems require a common understanding of the language of these systems. To support this common understanding: AI, autonomy, “in-the-loop”, and decision loops will be defined within the scope of this discussion. These definitions have been derived from commonly accepted definitions, however, they will be discussed within a military context.

Artificial Intelligence

The term artificial intelligence, also referred to as “AI”, has broad meaning in the social arena. It conjures ideas ranging from the relentless mechanized *Terminator T-800* killing machines, to the morally confused bodiless intellect of HAL 9000.[7] Within this discussion, the term AI will be defined as an engineered system(s) with the capability to independently: (1) perceive the environment; (2) understand the environment; (3) predict outcomes of events; and, (4) manipulate resources to effect a desired outcome. In addition to these four basic characteristics, AI may be capable of: re-assessing the rationality of their senses, knowledge, and frameworks, and modifying them to more accurately reflect reality. In short, AI systems may be able to adapt and learn.

This definition of AI is consistent other common descriptions. Russell and Norvig, two leading scientists in the field, describe AI as the ability to “perceive, understand, predict, and manipulate a world far larger and more complicated than itself.”[8] This definition, for brevity, has a noteworthy difference from the previous definition. By placing an intelligent agent, such as an AI, within the context of a larger system, Russell and Norvig emphasized that an AI’s knowledge and understanding is limited and ultimately subject to making mistakes. For instance, an AI system that has its sensors deceived, or its data “spoofed”, is as prone to making a faulty decision as a comparable human.

Summarizing, AI are engineered systems capable of performing several activities; perceiving, understanding, predicting (and choosing), manipulating (acting), and potentially learning from the environment in which they are immersed. Projecting these AI capabilities to the future, these systems will likely obtain advantages and proficiencies in many areas, such as manufacturing, vehicle operations (space, air, land, and sea), data/signal compilation and analysis, and simple rules-based tasks.[9] AI capabilities may still be unproven in the area of ill-defined problems, particularly those requiring value-based assessments on social, cultural, political or ethical issues.

For brevity, issues such as deep learning, behavior emulation, or programmed logic will not be explored. For the scope of this discussion, it is sufficient to assume that AI system designs will be wide and varied, with each method providing its own mix of capabilities, potential advantages, and vulnerabilities.

Autonomy

In the context of AI systems, there are *autonomous* and *semi-autonomous systems. Autonomous* is defined as having some significant degree of self-governance, granted by system design and/or system users; and, *semi-autonomous* is defined as a system having some low degree of self-governance. The degree of self-governance (autonomy) for AI systems can also vary among the internal activities [perceive, understand, predict (and choose), manipulate (act), and learn]. Degrees of autonomy may be a result of technical limitations or a consequence of varying levels of trust embedded in the AI system.

Preeminent futurists Parasuraman and Riley have examined and formally ascribed multiple hierarchies of autonomy.[10] However for the purpose of this discussion, a simple three level hierarchy is used; high, medium, and low. A brief introduction to the three levels of autonomy follows.

High autonomous AI systems are systems that operate and execute activities with minimal, if any, human oversight. From an operational point of view, high autonomous systems are those AI systems that have been vested with high levels of trust. These are systems that perform most of the five internal processes (perceive, understand, predict (and choose), manipulate (act), and learn) with minimal limitations and a high degree of independence.

Medium autonomous systems are systems that operate with some degree of oversight. From an operational point of view, medium autonomous systems are those AI systems that have been vested with a varying degree of trust for each of the AI activities. Operator oversight can be implemented in a variety of forms. Minimally restrictive oversight could be AI systems that need to notify a user of an action or intended action (passive oversight). This would allow end users to closely monitor and, if necessary, direct AI activities. These restrictions may be applied to specific internal processes (perceive, understand, predict (and choose), manipulate (act), and learn), resulting in a mix of autonomy within an AI system. For example, a moderately autonomous strategic planning AI system may be trusted with developing and comparing Course of Actions (COAs), but may not be fully trusted to select or execute a preferred COA. Oversight could be more restrictive, where authorization would be required before an activity could be pursued.

The differentiation between low (semi-autonomous) and medium autonomous systems is somewhat subjective. However, from an operational perspective, low degrees of autonomy implies a low degree of trust. Low autonomous AI systems are systems that are restricted in many of their internal processes (perceive, understand, predict (and choose), manipulate (act), and learn) and require positive authorization (active oversight). An example of this would be limiting an AI system from continued learning and self-altering programing. Corresponding trust in AI systems with modified and untested programming is likely to be low. As a result, emerging AI could have a low level of autonomy.

Autonomy may also be limited by technological capability or readiness reasons. Computing power and task complexity are examples where technological restrictions may occur. Similarly, autonomy can be intentionally restricted below a technological capability. AI restrictions may be the result of social or cultural tolerances for AI systems. Tolerance for machines making ethical decisions may result in limitations that are social (or cultural), rather than technologically, imposed.

From these discussions, a relationship for autonomy can be derived. Consider, autonomy directly proportional to the capability of the AI system, and the societal trust in those capabilities, and inversely proportional to the magnitude of ethical dilemma. This relationship can be presented as follows:

$$\text{Autonomy} \propto \frac{(\text{trust} \times \text{capability})}{\text{Ethical consequence}}$$

Figure 1. Autonomy as a Function of Trust, Capability and Ethical Consequence

Combining the definitions of AI and autonomy helps clarify an often confusing landscape of discussions about autonomous systems. For example, a projectile is not autonomous. A projectile does not perceive, understand, or predict; nor can it alter its point of impact. A projectile cannot be autonomous without first being intelligent. Likewise a traditional landmine is not an autonomous weapon. It cannot sense or discriminate between combatants and non-combatants; nor can it affect the outcome after it has been armed. A possible example of an autonomous weapon is an Automated Arial Defense System (AADS). It senses flying objects and discriminates between them. After identifying the flying threats, it can act by deploying counter measures to defeat them. Hence, an AADS meets the criteria of an AI system, by operating autonomously as it perceives (senses aircraft), understands (differentiates aircraft), predicts (assesses a high threat), and manipulates (deploys counter measures) without additional input.

In-the-Loop

The rising use of autonomous systems, such as UAVs and Remotely Piloted Vehicles (RPAs), to perform lethal strikes has prompted “in-the-loop” debates across the social and political spectrum. These debates often center on the appropriate level of autonomy for AI systems that perform actions with legal, moral, and/or ethical concerns. These debates are aimed at understanding and clarifying the connections between the actions of autonomous systems and the responsibility for those actions. For the purposes of this paper, there are three major categories of “*in-the-loop*” discussions:

1. *Human-in-the-loop* or (*man-in-the-loop*)
2. *Human-on-the-loop* or (*man-on-the-loop*)
3. *No-human-in-the-loop* or (*no-man-in-the-loop*)

These social discussions have taken two main forms, low level – semi-autonomous systems, commonly referred to as *human-in-the-loop* (often *man-in-the-loop*) systems and medium autonomy, commonly referred to as *human-on-the-loop* (often called *man-on-the-loop*) systems.

Human-in-the-loop refers to limits on AI system autonomy preventing actions without specific human action or authorization. These types of limits are often threshold driven. For example, if an action were to meet a legal, moral, or ethical threshold, such as an action that would reasonably result in the loss of life (or present a high degree of legal or ethical risk), human engagement and positive authorization could be required before an AI system could complete that action. An example of this would be the use of an RPA to attack a target with lethal force. In the case of an RPA, the aircraft operations are assisted by autonomous AI systems, but a human (pilot) controls the aircraft operations. Additionally, a human confirms the identification and selection of targets, and a human decides to employ lethal force. Responsibility for the intentional actions performed by an RPA, including the legal, ethical, and moral consequences, are clearly linked to the humans controlling the RPA.

Human-on-the-loop refers to medium level autonomous systems, where AI systems notify human operators of an imminent action, again along a legal, moral, or ethical threshold. However, positive authorization may not be required for the AI system to complete an action. An example of this would be an autonomous AI system identifying a target for lethal use of force: the AI system may be required to provide a human operator either, (1) an opportunity to prevent the use of lethal force, or (2) positive authorization for the use of lethal force. Under this scenario, the responsibilities for the intentional actions of an autonomous AI system become clouded. The engineers who designed the autonomous system, the operators of that system, and the autonomous system itself, may be considered as participants to the intentional actions taken.

The *no-humans-in-the-loop* form of autonomy provides a final category for consideration. These *no-human-in-the-loop* systems are systems where no human engagement, authorization, or notification, occurs where high autonomous AI systems operate in a military environment. The ability of AI systems to independently decide legal, moral, or ethical issues, such as the lethal use of force, raises an array of concerns and questions. This type of autonomy creates a murky causal link between the actions it performs and who, or what, is ultimately responsible for those actions.

These AI scenarios illustrate a dilemma of autonomy: unambiguous responsibility versus competitive advantage. An example of this responsibility-advantage dilemma can be examined through a further discussion of the AADS systems, where an effective decision window may be less than a few seconds. In this scenario, it may be impractical to wait for human intervention to assess whether the AI system has properly determined whether a target is benign (passage airliner in commercial air zone) or and existential threat (a cruise missile in military air space). A second example of a responsibility-advantage dilemma is the engagement of two autonomous systems. In this scenario a time-to-action competition emerges when two autonomous systems engage each other; and the “quickest” system wins. Assuming that *human-in-the-loop* systems respond slower, a competitive advantage exists for *human-on-the-loop* and *no-human-in-the-loop* systems. In these time sensitive competitive environments, where time can be the decisive element of victory, artificially limited autonomy may concede significant advantages to high autonomous AI systems. Nonetheless, maximizing the competitive advantage simultaneously increases the ambiguity of responsibility for the actions of those systems. High autonomous AI systems have no clear operator to directly ascribe responsibility. The responsibility resides within a labyrinth of systems and sub-systems created by an ensemble of anonymous system engineers.

Previous *human-in-the-loop*, *human-on-the-loop*, and *no-human-in-the-loop* autonomy discussions have generally focused on only one dimension of AI autonomy: engagement with the environment. The other elements of AI (perceiving the environment, understanding, and predicting the environment, and learning) and their relative autonomy are often neglected from consideration. Yet the degree of autonomy granted to these other AI elements are equally important. This is increasingly important to consider as the capability of AI expands. Notwithstanding, before discussing these elements further, it is important to examine what the *loop* is.

Decision Loops

Within the context of this discussion, the *loop* is a rational decision-making or problem solving process or cycle that produces an action. In many professional fields including business, litigation, and military strategy, significant research has been performed to formalize and optimize decision-making processes. The purpose of these formalizations is to understand, standardize, and replicate what are generally agreed to be good decision processes. This is regularly achieved by dissecting the decision process (or cycle when performed iteratively) into discreet steps that correspond to recognizable or distinct phases of a decision.

Some common decision loops include the: Military Decision-Making Process (MDMP) -- a seven-step process[11]; the Commander's Decision Cycle (CDC) –a four step process[12]; Deming's Plan-Do-Check-Act (PDCA) cycle[13]; and the Boyd cycle, commonly named after the four steps of the loop –Observe, Orient, Decide and Act ("OODA" or OODA Loop).[14] While there are many other systems that can: help maximize an approach; enhance financial gains; and create efficiencies in business; for the purposes of this study, these military decision systems will be only considered.

Amongst these military decision systems, the relative simplicity of the OODA Loop makes it an ideal tool to explore AI systems. Its intuitive four steps: observe, orient, decide, and act, are easily understood and closely align to the first four principle elements of AI as described by Russell and Norvig (perceive, understand, predict, and manipulate, as well as learn). The OODA Loop provides a clear and obvious framework to explore the implications of integrating AI systems across the spectrum of competitive military environments.

Implications Framework

An additional framework will be used to explore future AI systems. The framework focuses inwardly to explore how future AI may affect internal elements of the military force. The intent of these explorations is not to provide a comprehensive analysis of the implications of future AI systems. Rather, this model provides a starting point to begin exploring the implications of future AI systems, and the best ways to more fully integrate AI systems.

The Joint Capabilities Integration Development System (JCIDS) framework will be used to look at the internal implications to the military force. The JCIDS framework identifies key factors to consider when integrating activities or materiel into military operations. These JCIDS elements are Doctrine, Organization, Training, Materiel, Leadership, Personnel and Facilities; and are commonly referred to by the mnemonic DOTMLPF (pronounced Dot-Mil-P-F). This paper will explore or highlight possible impacts on military operations through the DOTMLPF lens.

Analysis

Increasingly, success in the battlespace has become about collecting information, evaluating that information, then making quick, decisive decisions. NCW, with its advanced command and control concepts, demonstrated this concept during the emerging phases of information age warfare. However, as the information age has matured, adversaries have adopted its core tenant -- winning in the decision space is winning in the battle space.[15] As a result, the competitive advantage that may have once existed has eroded. Additionally, the principal feature of information age warfare, the ability to gather, and store communication data, has begun to exceed human processing capabilities.[16] Maintaining a competitive advantage in the information age will require a new way of integrating an ever-increasing volume of data into a decision cycle.

Future AI systems offer the potential to continue maximizing the advantages of information superiority, while overcoming limits in human cognitive abilities. AI systems with their near endless and faultless memory, lack of emotional vestment, and potentially unbiased analyses, may continue to complement future military leaders with competitive cognitive advantages. However, these advantages may only emerge if AI is understood, properly utilized, and integrated into a seamless decision process.

The potential implications of future AI systems will be explored throughout the remainder of this discussion. The OODA Loop will be the overarching framework for this analysis. Using this framework, the purpose of evaluating AI systems will be to provide a methodical approach to explore: (1) how future autonomous AI systems may participate in the various elements of decision cycles; (2) what aspects of military operations may need to change to accommodate future AI systems; and, (3) how implementation of AI and its varying degrees of autonomy may create a competitive decision space.[17] The examination of potential implications of AI on internal military operations will also use the DOTMLPF framework.

Observe

This analysis begins with the first element of the OODA Loop, observe. The act of observation is far more than simply seeing. Observe implies two distinct but entwined activities. The first activity is the scanning the environment. The second activity is recognizing potentially significant or meaningful events. It is the combination of these two activities that forms the basis for a sentient observation. The automation of *observe* can be performed using AI systems, either as a singular activity or as part of a broader integrated analysis.

A good example of an AI observing comes from the automotive industry. "Lane departure warning" and "lane keeping" technologies are recent implementations of AI systems that are beginning to emerge.[18] In lane departure systems, a specific AI (a rules based, task oriented AI) system continuously senses the vehicle's position relative to its lane. Then, if the AI system identifies an abnormal condition, such as a vehicle no longer staying within its lane, it alerts the driver. This human-AI collaboration improves driver safety by giving the driver an additional "set of eyes," without absolving the human driver of responsibility to maintain control over the vehicle.

Observe-Autonomy

Systems that *observe* require sophisticated AI analyses and systems. And within these systems various degrees of autonomy can be applied. Further, because *observe* is a combination of different activities, the degree of autonomy for scanning the environment may differ from the degree of autonomy for recognizing potentially significant events. Likewise varying degrees of autonomy may be applied to very specific tasks integral to scanning and recognizing.

High autonomous AI systems may be allowed to select or alter scan patterns, times and frequencies, boundary conditions and numerous other parameters; potentially including the selection of the scanning platforms and their sensor packages. High autonomous AI systems, integrated into feedback systems, could also alter and potentially optimize the scanning process. Thus allowing AI systems to independently assess the effectiveness of previous scans and explore alternative scanning processes.

At the other end of the spectrum, low autonomous AI systems might be precluded from altering pattern recognition parameters or thresholds for flagging an event as potentially significant. In this domain, AI systems could perform potentially complex analyses, but with limited ability to explore alternative approaches to examine additional environmental data.

Observe-DOTMLPF

Fundamentally, the key output from the observe activity is the observation of data that creates a signal identifying a potential need for further analysis. While the clear separation between observing and orienting data might at first seem unique, many real world equivalents exist that demonstrate this separation. A military example of this would be the development of field reports that annotate the observation of hostile adversaries. The report developed by a field unit, following standard protocols and containing observational data about a potential sighting of an advisory force, is reflexively developed and passed up a chain of command. Organizationally, no further analysis is performed until some command or staff element decides to take action.

Extrapolating from this example, it is not difficult to infer that when AI systems operate as autonomous observation systems they could easily be integrated into existing doctrine, organizations, and training. However, there could be slight differences between AI systems and human observers. For instance, standards for land based sentry watch durations may need to be reconsidered in terms of AI system capabilities. This is especially true when we consider manned and unmanned mixed teams. For example: AI systems could operate with human security forces, each with potentially different endurance limitations. Sentry outpost locations and configurations described by existing Field Manuals may need to be revised to address additional differing considerations for AI systems, such as safety, degrees of autonomy, communication, and physical capabilities, and dimensions, and integration issues with human forces.

Taking a broader perspective, the potential for ubiquitous and ever-present autonomous AI observation platform presents a new dimension to informational security. The possibility of persistent, covert, and mobile autonomous observation systems, offer security challenges that we only have just begun to understand. Information security within the cyber domain is just one example of the emerging challenges that AI systems can create as they continue to influence the physical domain.

Orient

The second element of the OODA Loop is orient. In this context, orient is the processes and analyses that establish the relative significance and context of the signal or data observed. The OODA step “orient” correlates to the Russell and Norvig AI element *understand*. An observation in its original raw form is unprocessed data of potential interest. The orientation and prioritization of that observation, begins when the observation is placed the context of: (among other things) previous experiences, organizational, cultural, or historic frameworks; or other observations.

The degree of priority given to an observation then determines how resources will be committed to synthesize and incorporate the observation into an operational picture. If the priority is low, further analysis may not occur. If the priority is moderate, perhaps the observation is aggregated with other observations for later analyses, or rudimentary analysis. And if the priority is high, the observation may be evaluated in detail. Thus with limited resources and cognitive abilities, large fractions of observations may not be evaluated in detail or brought into the operational picture.

One of the principal challenges of today’s military leader is managing the ever-increasing flow of information available to them. The ease and low cost of collecting, storing, and communicating has resulted in a supply of data that exceeds the cognitive capacity most humans.[19] As a result, numerous approaches are being considered to maximize the capability of commanders to prioritize and develop data rich common operating pictures.[20] These approaches include improved graphics displays as well as virtual reality immersion systems. Each is designed to give a commander access to larger and still larger volumes of data. Still, if commanders are indeed saturated with information, further optimizing the presentation of too much data, may not significantly improve battlespace performance.

However, the emergence of AI systems capable of contextualizing data has already begun. The International Business Machine (IBM) Corporation has already fielded advanced cognitive systems capable of performing near human level complex analyses.[21] Moreover, it is expected this trend will continue and these types of AI systems will continue to displace humans performing many staff officer “white collar” activities.[22] However, up to this point much of the analyses performed by existing systems, such as identifying market trends, or evaluating insurance payouts, have been in environments with reasonably understood rules and boundaries.

Orient-Autonomy

Autonomy issues associated with AI systems orientating data and developing situational awareness pictures are complex. AI systems operating with a high autonomy can: independently prioritize data; add or remove data from an operational picture; possibly de-conflict contradictory data streams; change informational lenses; and, can set priorities and hierarchies. High autonomous AI systems could continuously ensure the operational picture is the best reflection of the current information. The tradeoff to this “*most accurate”* operational picture might be a rapidly evolving operational picture with little continuity that could possibly confound and confuse system users. This type AI might require a blind faith acceptance to the picture presented.

At the other end of the spectrum, low autonomous AI systems might not explore alternative interpretations of data. These systems may use only prescribed informational lenses, and data priorities established by system users or developers. The tradeoff for a stable and consistent operational picture might be one that is biased by the applied informational lenses and data applications. This type AI may just show us what we want to see.

Additional considerations arise concerning future human-AI collaborations. These collaborations may not be considered long-term relationships because every leader uses and prioritizes information differently. Generic AI systems that prioritize information based on a set of standard rules, may not provide the optimal human-AI paring. Instead, AI systems that are adapted to complement a specific leader's attributes may enhance his decision-making. These man-machine interfaces could be developed over an entire career. As such, there may be a need to ensure flexibility and portability in autonomous systems, to allow leaders to transition from job to job and retain access to AI systems that are "optimized" for their specific needs.

Orient-DOTMLPF

While this is not meant to be an exhaustive application of DOTMLPF, a few issues do stand out. The use of AI systems for the consolidation, prioritization, and framing of data may require a review of how military doctrine and policy guides the use of information. Similar to the development of rules of engagement, doctrine and policy present interesting challenges to developing rules of information framing. For example, doctrine or policy development could potentially prescribe or restrict the use of informational lenses. While applying a lens to organize information is not without precedent, under a paradigm where AI systems could implement doctrine and policy without question or moderation, the consequences of a policy change might create a host of unanticipated consequences.

Additionally, AI systems capable of consolidating, prioritizing, and evaluating large streams of data may invariably displace the staff that currently performs those activities.[23] This restructuring could preserve high level decision making positions, while vastly reducing personnel performing data compiling, logistics, accounting and a host of other decision support activities. The effect of this restructuring might be the loss of many positions that develop the judgment of future leaders. As a result, increased automation of data analytics, and subsequent decrease in staff supporting those activities, may create a shortage of leaders with experience in these analytical skills and tested judgment.

This brief assessment is not meant to provide a complete analysis of the impacts of AI. Numerous other considerations in each of the DOTMLPF elements should be explored in greater detail. However, some other obvious impacts also need to be considered. Such as: increased communication and power demands at forward and tactical operational centers; need for trained operators; and, materiel integration.

Decide

The third element of the OODA loop is "*decide"*. In this context, *decide* is the process used to develop, and then select a course of action, to achieve a desired end state. The OODA step "*decide"* roughly correlates the Russell and Norvig AI element "*predict*". As a rational action, a decision is the selection of a course of action that is forecasted to improve a situation. Prior to selecting a course of action, the military decision making process requires development of multiple courses of actions (COAs) and consideration of their likely outcomes. Then one can compare the outcomes of multiple COAs and select the COA with the preferred outcome.

While there are many techniques, models, and processes available to guide development and selection of a COA, they all share an understanding of the desired end-state, the means available, and the ways those means can be employed. The feasibility, acceptability, suitability, and risk of each COA can then be compared to enable an optimal COA to be selected or decided upon. The basis for developing COA's and choosing among them can be further categorized as: rules based, or values based, decisions.

Rules Based Decisions

From an overarching perspective, rules based decisions explore ends, ways, and means, through the lenses of feasibility and suitability, but do not actively address questions of acceptability or risk. Further, rules based decisions are prescriptive with the decision space closed. Examples of rules based decisions include: 1) selecting the shortest route to a destination: and, 2) a cashier dispensing correct change. In these examples there are multiple COAs or solutions to the problem. The COA's are bounded and well quantified. Because rules are absolute, there is no consideration of acceptability or risk in COA development or selection.

Rules are evaluation criteria based on abstraction of human values, in many cases numerous competing values. In the previous example of providing correct change, legal, and ethical values are intrinsic to the process of providing correct change. However, when applied as a rule, "provide correct change" contains no consideration (by a cashier) about the acceptability or risks of not providing correct change. The only consideration is compliance with the rule.

In this sense, if an AI system is using a rules based decision process, there is inherently a *human-in-the-loop*, regardless of the level of the AI autonomy. This is because human value judgments are inherently contained within the rule development process. From this perspective, when considering the concept of human *in/on-the-loop*, it would be worthwhile to include additional qualifiers of active or embedded human participation. Active *human-in-the-loop* implies an operator or external agent is assuming some responsibility for the value judgments. While AI systems with an embedded *human-in-the-loop* implies an engineered value judgment, thus anchoring responsibilities for the value judgments to the engineers of the AI system.

Value Based Decisions

Again, beginning from an overarching perspective, value based decisions explore ends, ways, and means, through the lenses of feasibility and suitability, while also potentially addressing issues of acceptability and/or risk. Value based decisions are generally associated with subjective value assessments, greater dimensionality and generally contain some legal, moral, or ethical qualities. The generation of COA's and their selection may involve substantially more nuanced judgments.

Additionally, differentiation of COAs may require evaluations of disparate value propositions. Values such as speed of an action, materiel costs, loss of life, liberty, suffering, morale, risk, and numerous other values often need to be weighed when selecting a COA for a complex issue. These subjective values, not easily quantified or universally weighted, can present significant challenges in assessing the level of autonomy to grant to AI decision activities. As automation continues to encroach into the decision space, these subjective areas may offer the best opportunities for humans to continue to contribute.

Decide- Autonomy

Autonomy for AI systems involved in the decision process can be divided into the two parts of the *decide* step, the development of COAs (and their predicted outcomes) and the selection of a preferred COA. The division of *decide* into two distinct activities may result in a mix in levels of autonomy used in the decision process. High levels of autonomy may be granted for the development of COA’s, while medium or low autonomy may be granted for the selection of a preferred COA. Alternatively, other combinations of high or low autonomy could be used.

High autonomous AI systems might be allowed to freely explore possible COAs, developing predictions about numerous possible end states. Low autonomous AI systems could also be designed to only explore COAs and predictions upon request. Likewise, higher levels of AI autonomy could be used to compare and select a preferred COA by using value-based decisions for highly capable systems, or rules-based systems for less capable systems. Alternatively, low autonomous AI systems might be only allowed to recommend COA’s for human selection.

Decide-DOTMLPF

While this is not meant to be an exhaustive application of DOTMLPF, a few issues do stand out with the use of AI systems in the decision process. The bifurcation of the *decide* step of the OODA Loop, COA development (and prediction), and COA selection, may apply to other decision cycle approaches such as the *Plan* step of the previously discussed Deming cycle. As a result, doctrinal discussions on decision cycles should be reassessed from an AI-human interoperability perspective. These examinations would ensure consistency between human decision cycle processes and those employed by the AI systems.

Additionally, the employment of value based or rules based decisions tends to vary according to the operational environmental and the level of operation. Tactical applications often tend towards rules based decisions, while operational and strategic applications tend towards values based decisions. Clarifying doctrine, training, and policies on rules based and values based decisions could be an essential element of ensuring that autonomous decision making AI systems are effectively understood, trusted, and utilized.

During the acquisition of AI systems, AI decision process categories (such as rules-based, values-based, emulation-based or other processes) should be understood and standardized. Thus creating clear categories of AI systems that may improve the system acceptance and the expectations by human operators and engineers. This could also aid in clarifying responsibilities for AI decisions between operators, systems, and engineers.

Other DOTMLPF considerations may include AI impact on leadership development opportunities. The ability of AI systems to contribute to planning processes may result in the reduction of planning staffs. These reductions could reduce the number of staff developmental opportunities for future leaders. Moreover, experience within those leadership pipelines may be diminished, if the roles and responsibilities are not designed to ensure development. Changing leadership training and assignment opportunities will likely have many more unintended consequences and requires additional study and analysis.

Act

The last element of the OODA Loop is *act*. The OODA step "*act*" closely correlates to the Russell and Norvig AI element *manipulate*. The "act" step converts the outcome of the OODA "*decide*" step and the complimentary Russell and Norvig "*predict (and chose)* AI element into an intentional event*."* For AI systems this ability to manipulate the environment may take several forms. The first form may be indirect, where an AI system concludes its manipulation step by notifying an operator of its recommendations. The second form may be through direct manipulation, both in the cyber and the physical or “real world” domains.

Manipulation in the cyber domain may include the retrieval or dissemination of information, the performance of analysis, the execution of cyber warfare activities, or any number of other cyber activities. In the physical realm, AI systems can manipulate the environment through mechanized systems tied into an electronic system. These mechanized systems may be a direct extension of the AI system or may be separate systems operated remotely.

Act-Autonomy

Within the OODA framework, once the decision has been made, the act is reflexive. For advanced AI systems, there is the potential for feedback to be provided and integrated as an action is taken. If the systems supporting the decision operate as expected, and events unfold as predicted, the importance of the degree of autonomy for the AI system (to act) may be trivial. However, if events unfold unexpectedly, the autonomy of an AI system to respond could be of great significance.

The authority and ability to alter an action may be viewed from multiple perspectives. One perspective suggests that an alteration of an approved action is only a minor change and in-scope of the original decision. From this perspective, any changes can be considered adjustments to the action, and not changes to the decision. A second perspective suggests that any alteration of an approved action, or deviation from expected conditions, is considered out-of-scope, voiding the approval of the original decision. Thus, in this lens, any further action would require a new complete iteration of the OODA Loop.

Consider a scenario where an observation point (OP) is being established. The decision to set up the OP was supported by many details. Amoung these concerns were: the path taken to set up the OP, the optimal location of the OP, the expected weather conditions, and the exact time the OP would be operational. Under a strict out-of-scope interpretation, if any of the real world details differed, even slightly, from those supporting the original decision, they would all be viewed as adjustments to the decision, and the decision would be voided. While under a less restrictive in-scope interpretation, if the details closely matched the expected conditions, they would be viewed as adjustments to the approved decision, and the decision would still be valid.

High autonomous AI systems could be allowed to make in-scope adjustments to the “act”. Allowing adjustments to the “act” would preclude a complete OODA cycle review. By avoiding this requirement, a new OODA cycle, an AI system might outperform low autonomous AI elements (and human oversight) and provide an advantage to the high autonomous system. Low autonomous AI systems following the out-of-scope perspective would be required to reinitiate a new decision cycle every time the real world did not exactly match expected conditions. While the extreme case may cause a perpetual initiation of OODA cycles, some adjustments could be made to the AI system to mitigate some of these concerns. The question still remains to determine the level of change that is significant enough to restart the OODA loop. Ultimately, designers of the system would need to consider how to resolve this issue.

Act-DOTMLPF

This is not a comprehensive examination of autonomous AI systems performing the *act* step of the OODA loop. Yet in the area of doctrine, training, and leadership an issue rises for quick discussion. Humans often employ assumptions when assigning/performing an action. There is a natural assumption that real world conditions will differ from those used in the planning and authorization process. When those differences appear large, a decision is re-evaluated. However, when the differences appear small, a new decision is not sought, and some risk is accepted. The amount of risk is often intuitively assessed and depending on personal preferences the action continues or is stopped. However because of the more literal nature of computational systems, autonomous systems may not have the ability to assess and accept “personal” risks. As a result military doctrine addressing command and leadership philosophies, such as *Mission Command* and decentralize operations, should be reviewed, and if necessary updated, to determine their applicability to operations in the information age.[24]

Conclusion

This paper addressed some of the challenges and gaps involving the integration of future AI systems into military operations. First, it provided some baseline definitions of key terms such as: artificial intelligence, autonomy, and degrees of autonomy, *in-the-loop,* and decision loops. Then, it considered AI systems performing four principle functions: perceive, understand, predict (and choose), and manipulate (act). These functions were then examined in respect to the decision model of the OODA Loop. The OODA Loop, with its four principle steps: observe, orient, decide, and act, closely aligned with the aforementioned four elements of AI systems and provided an approach to consider future AI systems for military operations.

This study also briefly explored some of the possible implications of AI systems using the Lens of DOTMLPF as another model for analysis. Through this lens it was demonstrated that the integration of future AI systems has the potential to permeate the entirety of military operations, from acquisition philosophies to human-AI team collaborations. Key issues identified in this study include a potential need to develop clear categories of AI systems and applications. These categories should be aligned along axes of trust, with rules-based and values-based decision processes clearly demarcated. This study proposed an equation to consider the relationship of autonomy and societal trust in AI systems. Because of the nature of machines to abide to literal interpretations of policy, rules, and guidance, a review of their development should be performed to minimize unforeseen consequences. This study established a coherent framework for future discussions about the integration of artificial intelligence systems in future military operations.

Endnotes

[1] George Galdorisi, "Keeping Humans in the Loop," *U.S Naval Institute Proceedings*, February 2015.

[2] General Philip Breedlove, quoted in Lolita Baldor, "Military Wants to Fly More Sophisticated Drones," *Fox News*, November 4, 2010, http://www.foxnews.com/us/2010/11/04/military-wants-fly-sophisticated-drones.html (accessed May 19, 2017).

[3] Alessandro Zocco and Lucio De Paolis, "Augmented Command and Control Table to Support Network-Centric Operations," *Defense Science Journal* 65 (January 1, 2015): 39-45.

[4] Carlo Kopp. "Understanding Network Centric Warfare," *Australian Aviation*, January/February 2005.

[5] Ibid.

[6] U.S. Joint Chiefs of Staff, J. C. I. D. S. *Manual for the Operation of the Joint Capabilities Integration and Development System* (Washington, DC: U.S. Joint Chiefs of Staff, 2012).

[7] The Terminator T-800 is the central character in the 1984 science-fiction movie titled *Terminator*. The movie follows the T-800 in its single-minded pursuit of its mission to kill its intended target. HAL 9000 is the artificial intelligence system piloting a space ship to the moons of Jupiter in the 1968 science-fiction movie titled *2001: A Space Odyssey*. A plot line of the movie follows HAL 9000 as it attempts to resolve an unexpected moral dilemma—should it keep secrets? HAL 9000 attempts to resolve the dilemma of secrecy by eliminating the humans in its care.

[8] Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Egnlewood Cliffs, NJ: Prentice-Hall, 1995).

[9] Aaron Smith and Janna Anderson. “AI, Robotics, and the Future of Jobs,” August 2014, http://www.pewinternet.org/2014/08/06/future-of-jobs/ (accessed March 5, 2017).

[10] Raja Parasuraman and Riley Victor. "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39, no. 2 (1997): 230-253.

[11] U.S. Department of Army, *Army Doctrine Publication 5-0: The Operations Process* (Washington, DC: U.S. Department of the Army, 2012).

[12] U.S. Joint Chiefs of Staff, *Joint Taskforce Headquarters*, Joint Publication 3-33 (Washington, DC: U.S. Joint Chiefs of Staff, 30 July 2012).

[13] Howard Gitlow et al., *Tools and Methods for the Improvement of Quality* (Boston: Irwin 1988), 159-160.

[14] Tim Grant, and Kooter Bas, "Comparing OODA & Other Models as Operational View C2 Architecture," *Proceedings of the 10th International Command and Control Research Technology Symposium*, 2005.

[15] Roger N. McDermott, *Russian Perspectives on Network-Centric Warfare: The Key Aim of Serdyukov's Reform* (Fort Leavenworth, KS: Foreign Military Studies Office (Army), 2011).

[16] Ang Yang, Abbass Hussein, and Sarker Ruhul. "Evolving Agents for Network Centric Warfare," *Proceedings of the 7th Annual Workshop on Genetic and Evolutionary Computation*, 2005, 193-195.

[17] Decision space is the range of options that military leaders explore in response to adversarial activities. Competitiveness in the decision space is based on abilities to develop more options, more effective options and to develop and execute them more quickly. Numerous approaches to managing decision space exist. NCW is an approach that emphasizes information rich communications and a high degree of decentralized decisions to generate options and “self synchronized” activities.

[18] BMW North America, “Lane Departure Warning,” http://www.bmw.com/com/en/insights/technology/technology_ guide/articles/lane_departure_warning.html (accessed March 5, 2017); American Honda Motor Company, “2016 Honda Pilot Lane Keeping Assist System,” http://hondanews.com/honda-corporate/channels/corporate-safety/videos/2016-honda-pilot-lane-keeping-assist-system (accessed March 5, 2017); Ford Motor Company, “Lane Keeping System,” https://owner.ford.com/how-tos/vehicle-features/safety/lane-keeping-system.html (accessed March 5, 2017).

[19] Yang, Hussein, and Ruhul. "Evolving Agents for Network Centric Warfare,” 193-195.

[20] Alessandro Zocco, and Lucio Tommaso De Paolis. "Augmented Command and Control Table to Support Network-centric Operations," *Defence Science Journal* 65, no. 1 (2015): 39-45.

[21] The IBM Corporation, specifically IBM Watson Analytics, has been employing “cognitive analytics” and natural language dialogue to perform “big data” analyses. IBM Watson Analytics has been employed in the medical, financial and insurance fields to perform human level analytics. These activities include reading medical journals to develop medical diagnosis and treatment plans; performing actuary reviews for insurance claims; and recommending financial customer engagement and personalized investment strategies.

[22] Smith and Anderson. “AI, Robotics, and the Future of Jobs”; Executive Office of the President, “Artificial Intelligence, Automation, and the Economy,” December 2016, https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files /documents/ Artificial-Intelligence-Automation-Economy.PDF,(accessed online 5 March 2017).

[23] Smith and Anderson. “AI, Robotics, and the Future of Jobs”; Executive Office of the President, “Artificial Intelligence, Automation, and the Economy,” December 2016, https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/ documents/Artificial-Intelligence-Automation-Economy.PDF, (accessed online 5 March 2017).

[24] Jim Storr, "A Command Philosophy for the Information Age: The Continuing Relevance of Mission Command," *Defence Studies* 3, no. 3 (2003): 119-129.