2022-05-02

Cybersecurity, Artificial Intelligence, and Risk Management: Understanding Their Implementation in Military Systems Acquisitions

Mun, Johnathan; Housel, Thomas

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/70374

SYM-AM-22-065

Excerpt from the Proceedings
of the
Nineteenth Annual Acquisition Research Symposium

Acquisition Research:
Creating Synergy for Informed Change

May 11–12, 2022

Published: May 2, 2022

The research presented in this report was supported by the Acquisition Research Program at the Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact any of the staff listed on the Acquisition Research Program website (www.acquisitionresearch.net).

Cybersecurity, Artificial Intelligence, and Risk Management: Understanding Their Implementation in Military Systems Acquisitions

Dr. Johnathan Mun—is a research professor at the Naval Postgraduate School. [jcmun@nps.com]

Dr. Thomas Housel—is a professor at the Naval Postgraduate School. [tjhousel@nps.edu]

Abstract

This research has the explicit goal of proposing a reusable, extensible, adaptable, and comprehensive advanced analytical modeling process to help the U.S. Navy in quantifying, modeling, valuing, and optimizing a set of nascent Artificial Intelligence and Machine Learning (AI/ML) applications in the aerospace, automotive and transportation industries and developing a framework with a hierarchy of functions by technology category and developing a unique-to-Navy-ship construct that, based on weighted criteria, scores the return on investment of developing naval AI/ML applications that enhance warfighting capabilities.

This current research proposes to create a business case for making strategic decisions under uncertainty. Specifically, we will look at a portfolio of nascent artificial intelligence and machine learning applications, both at the PEO-SHIPS and extensible to the Navy Fleet. This portfolio of options approach to business case justification will provide tools to allow decision-makers to decide on the optimal flexible options to implement and allocate in different types of artificial intelligence and machine learning applications, subject to budget constraints, across multiple types of ships.

The concept of the impact of innovative technology on productivity has applicability beyond the Department of Defense (DoD). Private industry can greatly benefit from the concepts and methodologies developed in this research to apply to the hiring and talent management of scientists, programmers, engineers, analysts, and senior executives in the workforce to increase innovation productivity.

Introduction

This research has the explicit goal of proposing a reusable, extensible, adaptable, and comprehensive advanced analytical modeling process to help the U.S. Navy in quantifying, modeling, valuing, and optimizing a set of nascent Artificial Intelligence and Machine Learning (AI/ML) applications in the aerospace, automotive and transportation industries and developing a framework with a hierarchy of functions by technology category and developing a unique-to-Navy-ship construct that, based on weighted criteria, scores the return on investment of developing naval AI/ML applications that enhance warfighting capabilities.

This current research proposes to create a business case for making strategic decisions under uncertainty. Specifically, we will look at a **portfolio of nascent artificial intelligence and machine learning applications**, both at the **PEO-SHIPS** and extensible to the **Navy Fleet**. This portfolio of options approach to business case justification will provide tools to allow decision-makers to decide on the **optimal flexible options to implement** and allocate in different types of artificial intelligence and machine learning applications, subject to budget constraints, across multiple types of ships.

The concept of the impact of innovative technology on productivity has applicability beyond the Department of Defense (DoD). Private industry can greatly benefit from the concepts and methodologies developed in this research to apply to the hiring and talent management of scientists, programmers, engineers, analysts, and senior executives in the workforce to increase innovation productivity.

Research Objective

The primary objective of the proposed research is to provide a business case analysis and ROI estimates for AI and ML systems and applications that will improve their acquisitions life cycle. Currently, the DoD has a portfolio of nascent artificial intelligence and machine learning applications, both at the PEO-SHIPS and eventually extensible to the entire Navy Fleet. The main research problem is to create business case examples on how this portfolio of AI/ML applications is valued and optimized. The portfolio of options approach provides business case justification, providing tools to allow decision-makers to down select the optimal flexible options to implement and allocate in different types of AI and ML applications, subject to budget constraints, across multiple types of ships.

Literature Survey

For the DoD, acquiring artificial intelligence (AI) technology is a relatively new difficulty (DoD). Given the significant danger of AI system acquisition failures, it's vital for the acquisition community to look at new analytical and decision-making methodologies for controlling these systems' acquisitions. Furthermore, many of these systems are housed in tiny, inexperienced system development firms, further complicating the acquisition process with insufficient data, information, and processes. The DoD's well-known challenge of obtaining information technology automation will almost certainly be compounded when it comes to acquiring complicated and dangerous AI systems. To assist in minimizing costly AI system acquisition disasters, more powerful and analytically driven acquisition approaches will be required. To complement existing earned value management, this study identifies, reviews, and proposes advanced analytically based methods of integrated risk management (Monte Carlo simulation, stochastic forecasting, portfolio optimization, and strategic flexibility options) and knowledge value-added (using market comparables to determine the economic value of intangibles and non-financial government programs).

The Real Options Valuation methodology is a new approach that has been effectively applied in a variety of commercial industries to measure the entire future worth of decisions taken when there is a significant degree of uncertainty at the time decisions are needed. PEO SHIPS needs a new methodology to assess the total future value of various combinations of nascent AI/ML applications and how they will enable affordable warfighting relevance over the full ship service life to successfully implement the Surface Navy's Flexible Ships concept.

This research project will look at how the Integrated Risk Management technique may be applied in the Future Surface Combatant Analysis of Alternatives to estimate the entire future value and return on investment of artificial intelligence design characteristics (AOA).

Defense Acquisition System

The Defense Procurement System, which supervises national investment in technologies, projects, and product support for the U.S. Armed Forces, handles the acquisition of new systems for the DoD (DoD, 2003). Its main goal is to "acquire high-quality goods that meet user objectives while delivering measurable advances in mission capability and operational support in a timely and cost-effective manner" (DoD, 2003). The Joint Capabilities Integration and Development System (JCIDS), the Planning, Programming, Budgeting, and Execution (PPBE) process, and the Defense Acquisition System are three different but interrelated processes inside the DoD Decision Support System (DoD, 2017). Within the Defense Acquisition System, this study focuses on program management rather than contract management.

ACATs are assigned to acquisition programs based on the type of program and the dollar amount spent or expected to be spent within the program (DoD, 2015a). Figure 1 depicts the Defense Acquisition System's numerous cost-based designations and categories. All ACAT classification dollar amounts are determined in fiscal year 2014 dollars (DoD, 2015a). ACAT I is for big defense acquisition programs with a Research, Development, Test & Evaluation (RDT&E) budget of more than $480 million, or a total procurement budget of more than $2.79 billion (DoD, 2015a). ACAT IA programs do not meet the criteria for ACAT I and will spend more than $835 million in total procurement (DoD, 2015a) or more than $185 million in RDT&E. ACAT II programs do not meet the criteria for ACAT I and will spend more than $520 million in total life-cycle cost, $165 million in the total program cost, or $40 million for any single year of a program (DoD, 2015a). Finally, ACAT III programs are those that do not meet the requirements for ACAT I or ACAT II (DoD, 2015a). Because each category has varied reporting requirements and designated decision-makers, the multiple designations allow for decentralized control of a program (DoD, 2017).

There are five phases within the Defense Acquisition System:

- Materiel Solution Analysis (MSA)
- Technology Maturation and Risk Reduction (TMRR)
- Engineering and Manufacturing Development (EMD)
- Production and Deployment (PD)
- Operations and Support (OS)

The acquisition process is driven by requirements for new or better capabilities, which are delivered through the JCIDS process (DoD, 2015a). The relationship between the acquisition and capabilities needs processes, as well as their interaction in the various acquisition phases, is depicted in Figure 2. The capabilities required from the JCIDS procedure are assumed to be correct and necessary in this investigation.

The Materiel Development Decision kicks off the MSA phase after an Initial Capabilities Document (ICD) has been validated (DoD, 2015a). Although an acquisition program is not legally constituted until Milestone B at the end of the phase, this choice kicks off the acquisition process (DoD, 2015a). The goals of the MSA phase are to select the most promising possible acquisition process solution that will meet the ICD's demands and to define the system's Key Performance Parameters (KPPs) and Key System Attributes (KSAs; DoD, 2015a). An Analysis of Alternatives (AoA) is used to assess the acceptability of proposed acquisitions based on "measures of effectiveness; important tradeoffs between cost and capacity; total life-cycle cost, including sustainment; timeline; the concept of operations; and overall risk" (DoD, 2015a, p. 17). During this stage, the PM is chosen and the Program Office is established (DoD, 2015a). After the necessary analysis is completed, the decision authority—usually the Defense Acquisition Executive (DAE), head of the DoD component, or Component Acquisition Executive (CAE), unless otherwise delegated—determines whether the program will proceed to the next phase based on the justification for the chosen solution, how affordable and feasible the solution is, and how adequate the cost, schedule, and other factors are (DoD, 2015a). Milestone A is the name given to this decision (DoD, 2015a). The MSA phase examines all possible solutions to a stated demand and, as a result, may be an opportune time to investigate strategic techniques like KVA or IRM.

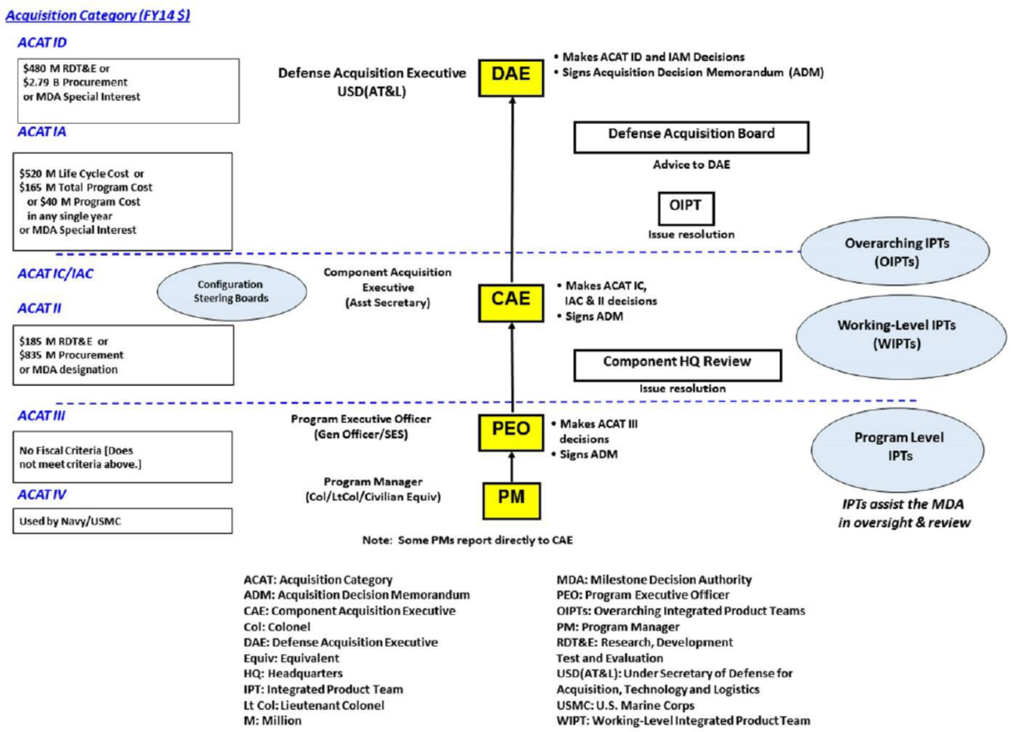


Figure 1. Acquisition Categories
(DoD, 2017)

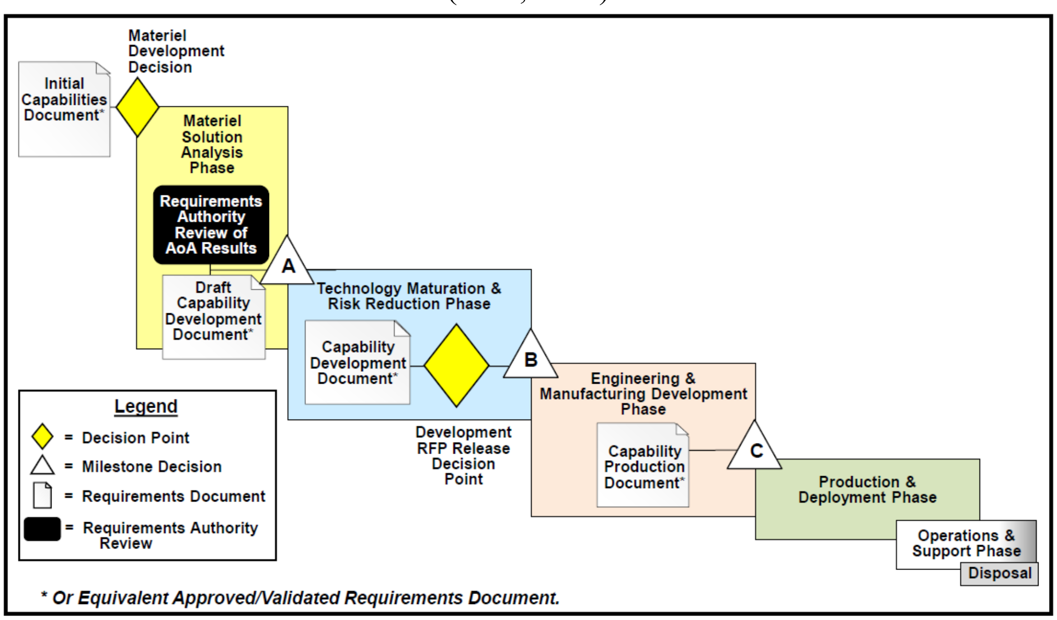


Figure 2. Interaction of Capabilities Requirements and Acquisition Process
(DoD, 2015a)

The program enters the TMRR phase after Milestone A approval to decrease the risk associated with the technology, engineering, life-cycle cost, and integration of the program before moving on to the EMD phase (DoD, 2015a). At this step, design and requirement trades are carried out based on the budget, timetable, and possibility of completion (DoD, 2015a). Contractors prepare early designs, including competing prototypes if practicable within the program, to show the practicality of their proposed solutions to the program office, guided by the acquisition strategy authorized at Milestone A (DoD, 2015a).

Technology Readiness Levels (TRLs) are a set of standards that show the level of risk involved with a solution maturing on time (DoD, 2015a). Technology Readiness Assessments (TRAs) are a metric-based technique for assessing the maturity and risk associated with important technology in an acquisition program (DoD, 2011). Each important technology in a program will be assigned a TRL by a TRA, ranging from 1 to 9 from lowest to maximum readiness level (DoD, 2011). Additional tools, such as IRM, to estimate the chances of a program remaining on schedule and on budget, may be useful at this stage. The Publication Decision Point for Development Requests for Proposals (RFP) permits the release of an RFP with firm and clearly specified program requirements for contractors to submit bids (DoD, 2015a). Unless the milestone decision authority waives it, the Preliminary Design Review (PDR) occurs prior to the completion of the TMRR phase (DoD, 2015a). Milestone B approves a program's entry into the EMD phase, awards a contract, and establishes the Acquisition Program Baseline (APB; DoD, 2015a). The APB is a legal commitment to the milestone decision authority that outlines the authorized program, especially the cost and schedule over the program's life (DoD, 2015a).

Once Milestone B has been approved, EMD can commence. Prior to production, the material solution is conceived, produced, and tested to ensure that all requirements have been met (DoD, 2015a). The hardware and software designs have been finished, and prototypes have been developed to detect any design flaws that will be uncovered during developmental and operational testing (DoD, 2015a). Federal regulation requires DoD procurement projects with a contract value higher than $20 million to utilize EVM to track and report program progress, which begins during this phase (DoD, 2019a). The manufacturing or software sustainment methods, as well as production capabilities, must be appropriately proven once a stable design that meets the given requirements have been validated (DoD, 2015a). Milestone C verifies that these requirements have been met and authorizes the start of the PD phase (DoD, 2015a).

The goal of the PD phase is to deliver a product that meets the standards established earlier in the process (DoD, 2015a). Low Rate Initial Production (LRIP) for manufactured systems or limited deployment for more software-intensive programs occurs first, with the system undergoing Operational Test[ing] and Evaluation (OT&E) to verify that stated criteria were satisfied (DoD, 2015a). Full-rate manufacturing occurs when the fielded systems have been approved and the product is deployed to operating units (DoD, 2015a). At this time, design changes are limited; however, some may still be made in response to identified flaws (Housel et al., 2019a). During this phase, contracts often revert to a fixed pricing strategy, lessening the PM's focus on cost and schedule variance (Housel et al., 2019b).

The operating system is meant to keep the product supported and performing well throughout its life cycle, which ends with the system's disposal (DoD, 2015a). Because operational units are using the product while production is ongoing, the OS phase overlaps with the PD phase, starting after the production or deployment decision (DoD, 2015a). PMs will maintain the system running by following the Life Cycle Sustainment Plan (LCSP) set during the purchase phase and providing the appropriate resources and support (DoD, 2015a). Technological upgrades, modifications due to operational needs, process enhancements, and other activities that may necessitate LCSP updates are all examples of sustainment and support (DoD, 2015a).

PMs employ six different models to develop their program structure, four of which are standard and two of which are hybrid, depending on the type of system being purchased (DoD, 2015a). These standard models serve as templates for hardware-intensive projects, defense-specific software-intensive programs, software-intensive programs that are incrementally deployed, and expedited acquisition programs (DoD, 2015a). The hybrid models, as seen in

Figure 3, combine the progressive character of software development with a hardware-centric program. Before attaining the Initial Operating Capacity, software development is arranged through a sequence of tested software builds that will climax with the completely required capability (IOC; DoD, 2015a). The incremental builds are timed to coincide with prototype hardware testing and other developmental requirements (DoD, 2015a). With the exception of the accelerated program, all other models use the same basic foundation across the five phases. AI and IT systems, as well as their connections to weapon systems, facilities, and Command, Control, Communications, Intelligence, Surveillance, and Reconnaissance (C4ISR), are becoming more common within the DoD (C4ISR; DoD, 2015b). As a result of the integration, enemies pose a greater security risk, emphasizing the significance of good cybersecurity skills and processes (DoD, 2015b). The DoD manages cybersecurity policy using the Risk Management Framework (RMF), which employs security measures based on risk assessments throughout a system's life cycle (DoD, 2015b). "All DoD IT that receives, processes, stores, displays, or transmits DoD information" is covered by the RMF (DoD, 2014, p. 2). The RMF's definition of cybersecurity goes beyond information security to include things like stable and secure engineering designs, training and awareness for all program users, maintainers, and operators, and the response, recovery, and restoration of a system after an internal or external failure or attack (DoD, 2015b). Figure 4 depicts the six steps of the RMF's procedure, which occurs throughout the acquisition process. The first stage is to categorize the system, which includes assessing the possible impact of a breach and describing the system and its boundaries (DoD, 2014). The RMF team is formed, the security plan is implemented, and the system is registered with the DoD Component Cybersecurity Program (DoD, 2014). The ICD includes cybersecurity standards, which drive MSA concerns during the AoA phase (DoD, 2015b). A cybersecurity breach might have serious consequences for missions, according to the risk assessment (DoD, 2015b). The RMF provides a somewhat objective technique for determining the cybersecurity risk level, as well as the basic baseline security controls that must be incorporated in the system's purchase strategy (DoD, 2015b).

The RMF team determines security measures in step two, including those that are common to other DoD programs (DoD, 2014). A plan is designed and recorded for regularly monitoring the effectiveness of the controls (DoD, 2014). The security plan is subsequently submitted to the DoD Components, who examine and approve it (DoD, 2014). During the MSA phase, the acquisition and cybersecurity teams collaborate to ensure that the proper level of security is applied throughout the program's life cycle, as well as in the system architecture and design (DoD, 2015b). During the MSA, the continuous monitoring strategy and security plan are also designed (DoD, 2015b).

The approved security procedures are then implemented in accordance with DoD specifications (DoD, 2014). The implementation must be well documented in the security plan for the system (DoD, 2014). In the TMRR phase, cybersecurity requirements are included in the system performance requirements (DoD, 2015b).

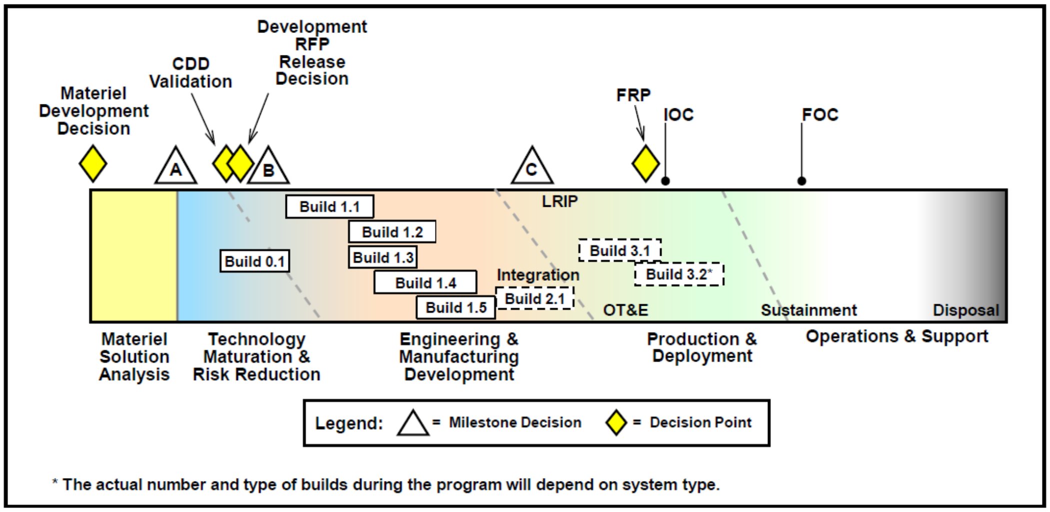


Figure 3. Hardware-Dominant Hybrid Program
(DoD, 2015a)

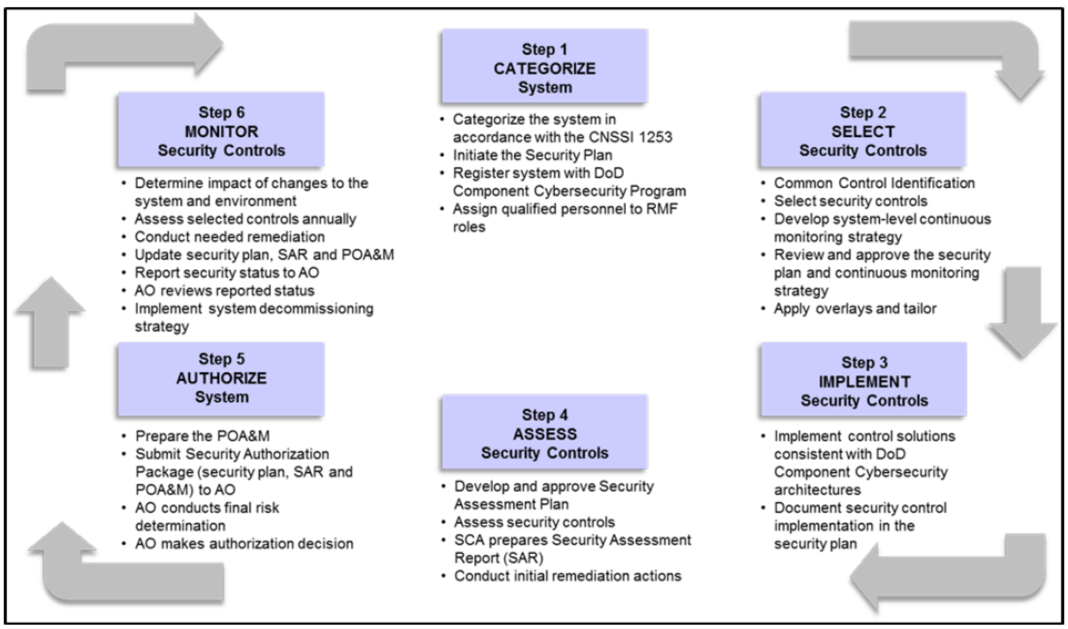


Figure 4. Risk Management Framework Process
(DoD, 2014)

The RMF team must then create, review, and approve a Security Assessment Plan that will allow the security controls to be properly assessed (DoD, 2014). Following approval, the security of the system is evaluated in line with DoD assessment processes and the Security Assessment Plan, during which vulnerabilities are assigned severity levels and the security risk for both the controls and the whole system is established (DoD, 2014). This is documented in the Security Assessment Report, which is necessary before any system is authorized, and security control repair activities are carried out (DoD, 2014). Prior to issuing an RFP, the Capability Development Document's cybersecurity criteria are evaluated throughout the TMRR process (DoD, 2015b). The cybersecurity parts of the Preliminary Design Review, which is also done during the TMRR process, will ensure that the authorized plan is executed in the chosen design and risks are reduced to an appropriate level (DoD, 2015b). All computer code follows applicable standards and secure coding practices as the system grows in the EMD phase, with evaluations undertaken and documented in the Security Plan (DoD, 2015b).

A Plan of Action and Milestones (POA&M) is produced based on the identified vulnerabilities, which identifies activities to mitigate the vulnerabilities, resources required to fulfill the plan, and milestones for completing tasks (DoD, 2014). The Security Authorization Package is given to the Authorizing Official who will decide whether the risk level is appropriate before authorizing the system (DoD, 2014). The POA&M is created during the MSA phase and continues throughout the system development process (DoD, 2015b).

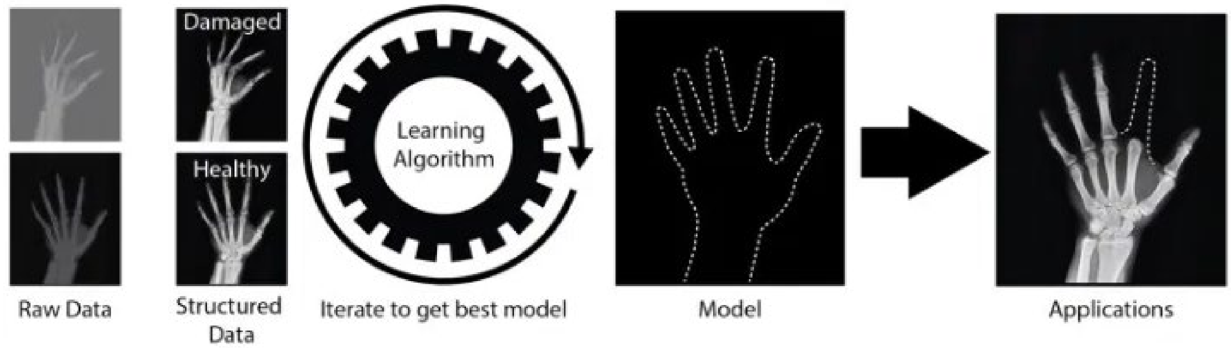
Finally, security controls must be monitored throughout the system's life cycle to ensure that any changes to the system or environment do not compromise cybersecurity (DoD, 2014). If vulnerabilities are discovered, the necessary remedy will be carried out, and the security strategy will be updated (DoD, 2014). The cybersecurity of a system is monitored in line with the continuous monitoring strategy and Security Plan once it has been approved and operationally implemented (DoD, 2015b). When the system, its surroundings, or the anticipated use of the system change, new risk assessments are done (DoD, 2015b). If a vulnerability is discovered, the PM changes the Security Plan and the POA&M to specify how the issue will be resolved (DoD, 2015b).

State of the AI

Machine Learning

Intelligence is the ability to process a specific sort of data, allowing a processor to solve significant problems (Gardner, 1993). Beyond the traditional idea of a person's analytic intelligence quotient (IQ), which can sometimes evaluate merely how well someone performs on an IQ test rather than their natural talents, psychologists have postulated many categories of intelligence. Howard Gardner (2003) proposed a theory of multiple intelligence, which suggests that traditional psychometric views of intelligence are too narrow and that intelligence should be expanded to include more categories in which certain processors, in this case, people, are better at making sense of different stimuli than others. Visual-spatial, linguistic-verbal, interpersonal, intrapersonal, logical-mathematical, musical, body-kinesthetic, and naturalistic intelligence are some of the categories of intelligence (Gardner, 1993). A counter-argument would be that these categories simply represent learned and disciplined habits that people develop through time as a result of their personality and environment. Regardless, both definitions of intelligence (traditional and many) are relevant to the stages involved in developing an artificial intelligence machine.

A computer can execute computations depending on the input data and produce an a priori defined outcome. It can be built and reprogrammed to repeat particular stages or algorithms and even change its conclusions based on previously calculated results using error-correcting techniques. The underlying principle of machine learning is a combination of these two phases. A computer system is fed data that is structured in such a way that the algorithm can identify it, deduce patterns from it, and make assumptions about any unstructured data that is presented later (Greenfield, 2019). In an x-ray learning algorithm, this is shown in Figure 5.



The image shows the steps an AI algorithm goes through in order to make a recommendation to a physician on where a missing body part should be. It takes in structured data and develops its understanding of what "right" looks like. When given unstructured data, it compares the image against previously trained models and identifies the abnormality with a recommendation on where to apply a fix, such as a prosthetic.

Figure 5. AI Training Algorithm
(Greenfield, 2019)

The basic concept of machine learning is illustrated in Figure 5, although the current research focuses on the many types of learning from the standpoint of procurement. The following are interpretations of different forms of learning in procurement algorithms provided by Sievo (2019), an AI procurement software business.

Supervised Learning

The patterns are taught to an algorithm using previous data, and it then recognizes them automatically in new data. Humans give supervision in the form of the right responses, which train the algorithm to look for patterns in data. This is a term that is widely used in procurement sectors like spend classification (Sievo, 2019).

Unsupervised Learning

The algorithm is set up to look for novel and fascinating patterns in brand-new data. The algorithm isn't expected to surface specific accurate answers without supervision; instead, it hunts for logical patterns in raw data. Within important procurement functions, this is rarely employed (Sievo, 2019).

Reinforcement Learning

The algorithm determines how to act in specific scenarios, and the behavior is rewarded or punished based on the outcomes. In the context of procurement, this is mostly theoretical (Sievo, 2019).

Deep Learning

Artificial neural networks gradually develop their capacity to accomplish a task in this sophisticated class of machine learning inspired by the human brain. This is a new opportunity in the procurement world (Sievo, 2019).

Natural Language Processing

Anyone who has used devices that appear to be able to understand and act on written or spoken words, such as translation apps or personal assistants like Amazon's Alexa, is already familiar with NLP-enabled AI. NLP is a set of algorithms for interpreting, transforming, and generating human language in a way that people can understand (Sammalkorpi & Teppala, 2019). Speech soundwaves are converted into computer code that the algorithms understand. The code then translates that meaning into a human-readable, precise response that can be applied to normal human cognition. This is performed by semantic parsing, which maps the language of a passage to categorize each word and forms associations using machine learning to represent not just the definition of the word, but also its meaning in context (Raghaven & Mooney, 2013). Figure 6 depicts this categorization and analysis process in the context of a procurement contract.

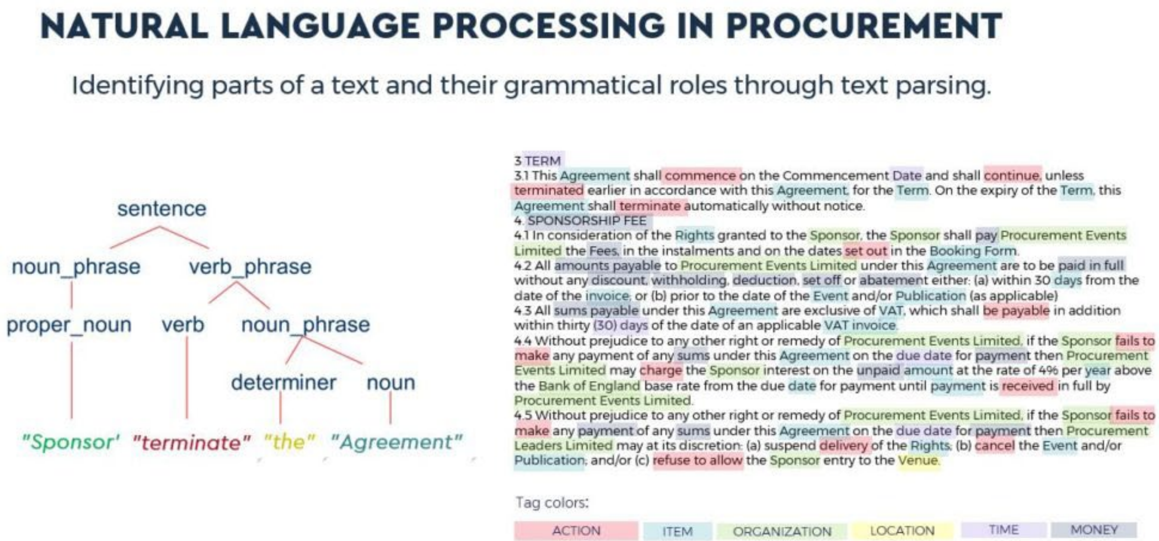


Figure 6. Semantic Parsing in Procurement
(Sievo, 2019)

Robotic Process Automation

Robotic Process Automation (RPA) is not AI; rather, it is an existing process that has been advanced by AI, as explained in the third section of this paper. RPA is defined as “the use of technology by employees in a firm to set up computer software or a robot to capture and interpret current applications for processing transactions, altering data, triggering reactions, and communicating with other digital systems” (Institute for Robotic Process Automation & Artificial Intelligence [IRPA & AI], 2019). When used correctly, robotic automation offers numerous benefits because it is not constrained by human limitations such as weariness, morale, discipline, or survival requirements. Robots, unlike their human creators, have no ambitions. Working harder will not get you more money or get you promoted, and being permanently turned off will have no effect because robotic automation just duplicates the practical parts of the human intellect, not the underlying nature of mankind (Zarkadakis, 2019). (Note, however, that machine learning relies on an incentive system to make judgments about positive or negative reactions.)

A future AI-enabled RPA option is for a machine to learn how to control the source of positive reinforcement fully independent of the rules required to achieve its aim. Things that survive develop to do so because of positive reinforcement from their environment and the fact that they continue to act in a way that is considered survivable. This should be taken into account in any future AI efforts, and especially in the case of why a human must always be present when final judgments are made. Regardless of whether AI systems have a perfect track record or not, they should not be entirely trusted.

Technology Trust

The Turing Test was created to test the capabilities of AI, as detailed in the third section of this report. Google developers designed Duplex, a spoken-word NLP tool, in 2018 to interface with its AI assistant. Its goal is to make phone calls on behalf of humans, converse with other humans, and respond to inquiries in a natural manner, all while sounding human (Leviathan, 2018). The algorithm can search for the information required as if it were a human searching for it on Google, for example. The AI assistant then calls a restaurant, for example, to schedule an appointment with the assistant's human. After being given oral information from a person hearing the orders, the software stutters, pauses, and elongates certain vowels as though it has to think about what it is saying, and responds with other recommendations within its limitations.

The authors questioned an AI NLP program named 1558M about one of the research issues twice for the purposes of this paper, and the machine responded with an unusual "opinion" of a negative and cautionary character (Figure 7). This tool was built to allow users to experiment with Open AI's new machine learning model (King, 2019). What's noteworthy about these responses is that they're all original, which means a search of the phrases turns up no copies. However, the language and tone make them sound like they're coming from a knowledgeable source, with just enough small evidence on the topic to be credible. However, the program does not finish its last phrase, making it unfinished but noteworthy. Clearly, such AI capabilities have a lot of potential for helping someone integrate with the DoD.

Is DoD ready for AI in contracting?

The answer is, "No." The DoD has an extremely limited ability to respond to technological shifts in a timely manner. In addition to the problems inherent in implementing a new technology on the same scale as DoD systems, there are several reasons to be cautious with new contracting methods:

The DoD does not have sufficient funds and resources to do so.

Contractors are generally not the most skilled in applying modern technology.

Contractors generally have limited autonomy.

A system cannot be fully tested within a period of one year, as such it would require extensive testing to determine whether or not the system is a success or failure.

As a result, it is unlikely that the DoD will adopt automation, AI, or robotics without the assistance of other agencies. For example, DoD is developing a framework for "Autonomy-Enhanced Contracts," which will "ensure that the DoD's contracting activities will be more

Is DoD ready for AI in contracting?

The Defense Science Board recently issued a report calling for AI to be an essential part of all military contracting. However, it's important to note that this is a new concept that hasn't been fully explored.

The report is largely based on research into military operations conducted in Afghanistan by the U.S. Army and the U.S. Navy. The report's recommendations include developing AI that can work in concert with humans to augment operational capabilities. AI systems could assist with the creation of better weapons systems, or assist human commanders to do their jobs more effectively.

The report also recognizes that AI is not a silver bullet: It's not enough to simply give a robot an A.I. license. Rather, it's important to understand that it's a "toolbox," and that AI could play an important role in the overall mission. It's the human factor that will be key to ensuring that AI systems are reliable and effective. The report

Figure 7. Two Separate Results from an AI Called 1558M
(King, 2019)

Explainable Reasoning

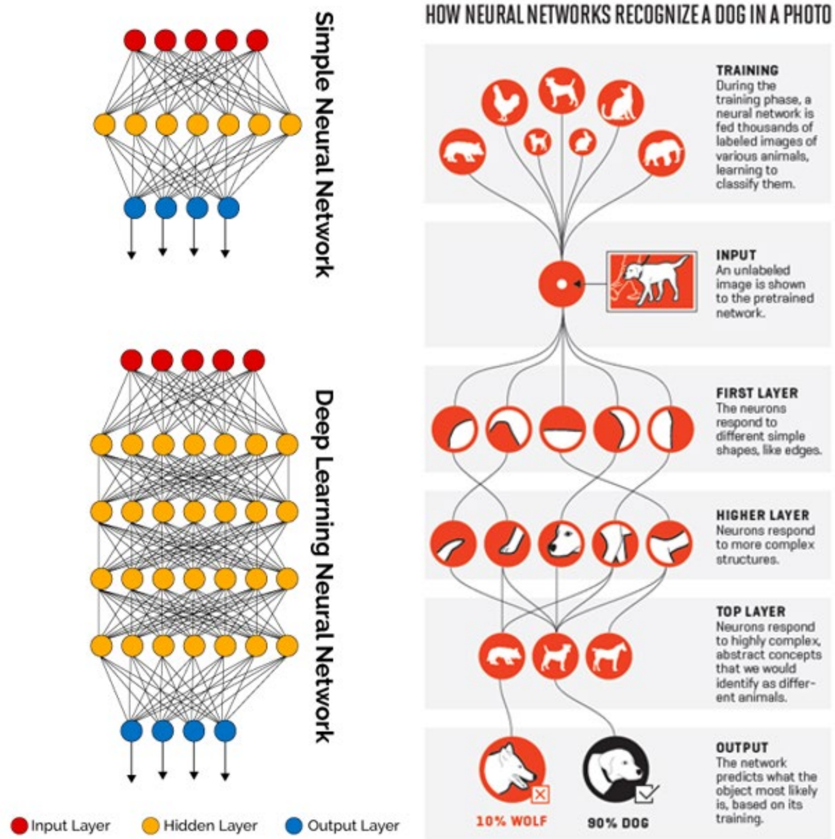
One of the barriers to AI adoption is the ability to explain how the algorithm arrived at its conclusions, which is necessary for auditing (Knight, 2017). It would be irresponsible to utilize artificial intelligence for military or financial goals without the capacity to track how judgments were made. Figure 8 depicts how AI currently categorizes data. The AI programs that produce the required outcome come up with their own means of navigating its layer complexities to develop output for the plethora of training data that went into creating the program.

Fortunately for the DoD, the Defense Advanced Research Projects Agency (DARPA), which is already ingrained in the defense ecosystem, is leading the charge on explainable AI research (Gunning, 2017). DARPA

> has taken the lead in pioneering research to develop the next generation of AI algorithms, which will transform computers from tools into problem-solving partners. DARPA research aims to enable AI systems to explain their actions, and to acquire and reason with common sense knowledge. DARPA R&D produced the first AI successes, such as expert systems and search, and more recently has advanced machine learning tools and hardware. DARPA is now creating the next

wave of AI technologies that will enable the United States to maintain its technological edge in this critical area. (DARPA, 2019)

The mechanics of how a Deep Neural Network navigates its trained data to identify different photographs are shown in Figure 9. Photos can be used to train an AI software, and associations of these trained data can then be used in the neural network to classify an input and eventually reach a conclusion. As a result, if the DoD decided to pursue human-machine cooperation in areas like contracting, its organic system would enable it to do so.



To identify the output layer, the Simple Neural Network uses a set of input data that only passes through one hidden layer. To better identify the output data, the Deep Learning Neural Network transmits the input data through numerous layers. The Deep Learning Neural Network goes through simple to more detailed layers of trained data that correspond with dog features to make a 90% confidence classification that the picture is a dog and a 10% possibility that it is a wolf to classify input data to determine if the given picture is a dog.

Figure 8. Simple Neural Network Compared to Deep Learning Network (Golstein, 2018; Parloff, 2016)

Human-Machine Partnership

Because sensor, information, and communication technologies generate data at rates faster than people can digest, comprehend, and act on, DARPA believes AI integration is vital as a human-machine symbiosis (DARPA, 2019). Machines are better at certain things, as they were throughout the industrial revolution, and using machines for those activities frees humans to become more productive in other areas. Separate areas of processing are where humans and machines flourish. Consider the following contrasts between computers and humans: calculate vs. decide; compare vs. make judgments; apply logic vs. empathize; unaffected by tiresome repetition vs. preferences; deals with enormous data vs. intuitional concentration on the most important (Darken, 2019). And while AI is capable of performing some jobs on its own, it performs better when paired with a human partner. Without sufficient restrictions, AI is a trusting learning system that can be manipulated by evil actors. According to certain studies, AI can be misled in ways that humans cannot owing to human intuition. Another study has been able to deceive a self-driving car into thinking a benignly tampered-with stop sign was a speed limit sign (Figure 10), which would almost certainly result in collisions if the car was left unattended (Eykholt et al., 2018).

Many people are aware of contemporary intelligent machine relationships that they may encounter on a regular basis without even realizing it. Google is the most popular search engine on the Internet because it gives more user happiness than its competitors, as stated with its other apps (Shaw, 2019). Google is so widely used as the primary search engine that many refer to it as “Googling” while looking for something online. This is a good example of humans engaging organically with a Bidirectional Encoder Representation-based AI system (BERT; Nayak, 2019). This is a strategy that trains a machine to answer a user’s inquiry based on the meaning of the words in the context of the question rather than on individual phrases. For example, when asking what time it is right before lunch, the user is really asking when they can eat; the outright answer would give the actual time, and the asker would deduce eating time, which was the underlying meaning of the question; the outright answer would give the actual time, and the asker would deduce eating time, which was the underlying meaning of the question. Another example of human contact with intelligent machines is so-called self-driving autos. The user mostly sits in a supervisory role while the automobile takes over one of the most dangerous moments in their lives and handles all road tasks autonomously to drive (Darken, 2019).

Contractors that rely on an AI system to make all of their decisions are vulnerable to deliberate misdirection by adversaries providing hostile information for competitive advantage or disruption. Fraudsters can learn how to manipulate computer algorithms, but only humans can assess the outcomes. AI software, on the other hand, can quickly extract data and explain contract content. It can swiftly gather and organize renewal dates and terms from a large number of contracts. It can help businesses evaluate contracts faster, organize and locate vast amounts of contract data more readily, reduce the risk of contract disputes and adversarial contract negotiations, and improve the number of contracts they can negotiate and execute (Rich, 2018).

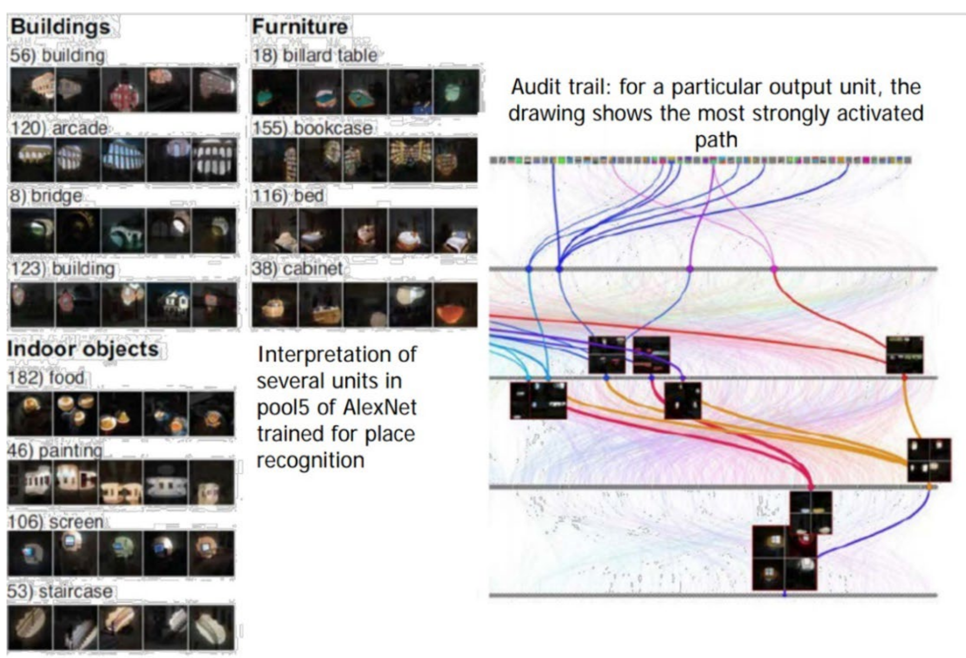
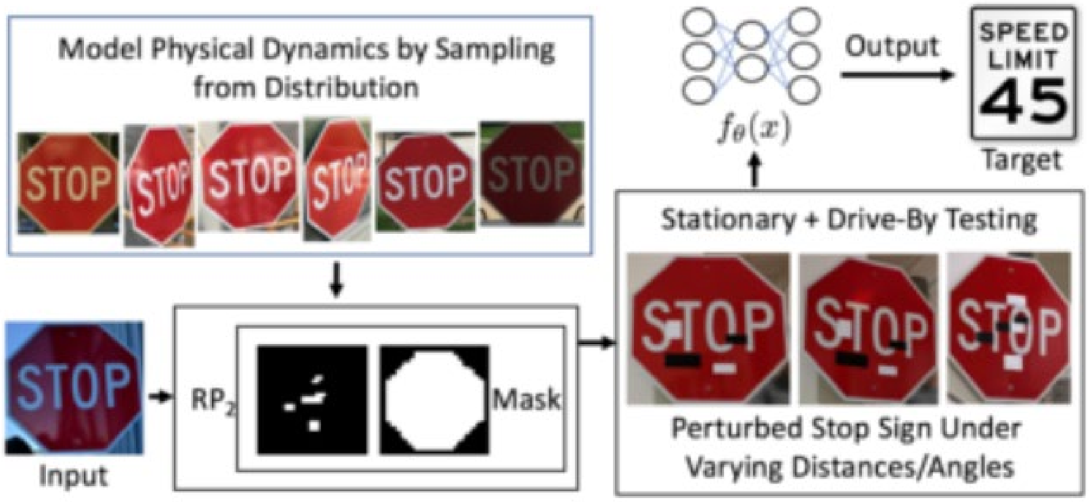


Figure 9. Visualization of Explainable AI
(DARPA, 2019)



An AI program in a self-driving car has trained data about a stop sign in its algorithm. When a target sign is seen in its environment, it references the trained data. As a test, researchers attached benign interruption markers on the sign, which confused the AI program to think the stop sign was a speed limit sign.

Figure 10. AI System Interpreting a Stop Sign
(Eykholt et al., 2018)

Case Study of Private Sector AI Application to Contracting

To compare DoD procurement options, we look at analogous situations in the private sector in the United States. Lawgeex is an example of a startup that is integrating AI into the procurement process in the private sector. An example contract component, the Non-Disclosure Agreement (NDA), demonstrated that AI software could outperform U.S.-trained lawyers with an average accuracy of 94%, compared to 85% for humans (Lawgeex, 2018). Large firms that rely on contracts to engage with partners, suppliers, and vendors have an 83% dissatisfaction rate with their organization's contracting processes, according to the report (Lawgeex, 2018). Another example is Icertis, which provides services to huge and well-known firms like 3M, Johnson & Johnson, and Microsoft, to name a few (Icertis, n.d.-a). Icertis offers a cloud-based AI platform that learns from the client's contracts, as well as control measures, to generate and help in contract setup, contract operations, governance, risk, and compliance, and reporting (Icertis, n.d.-a).

The fact that business is more acclimated to putting professional papers on digitally accessible storage infrastructure, whether local hard drives or the cloud, makes this practical now, rather than when it was initially theorized decades ago (Betts & Jaep, 2017). Nontechnical barriers to a completely automated contract review and analysis process now exist, such as the gathering of contract performance data, the disclosure of private contracts and their associated performance data, and changes in ethical limits on computer usage in legal practice (Betts & Jaep, 2017). The authors of these barriers also propose policy solutions to address them: begin using contract management software as a forcing function to create data in an AI teachable format, expand copyright protection for vendors to protect their intellectual property, and develop new rules to help mitigate AI risks so that it can work (Betts & Jaep, 2017).

Cloud-Based AI

We look at the concept of cloud computing to understand how AI may be disseminated throughout a system, update regulations, and learn from various human teachers in real-time. When it comes to DoD technology adoption, the term "speed of relevance" is frequently used.

The term “cloud” is used in the 2018 DoD Cloud Strategy to refer to an offsite physical IT infrastructure. This external infrastructure connects to a user’s PC through the Internet to access data servers that store information and run centrally managed operating systems like Microsoft Windows. This means that every user has the same software computing capacity and access to the most recent software, regardless of their organization’s IT professional talent or software budget. Organizations can have as much or as little access to what they need for projects as they need it, and they are unaffected by surges in demand or periods of inactivity, which now add to the cost of DoD systems (Shanahan, 2018). The DoD’s goal is to have AI-assisted rapid decision-making in a secure and visible data environment for increased operational efficiency.

> Data stored in an enterprise DoD cloud will be highly available, well-governed, and secure. Data will be the fuel that powers those advanced technologies, such as ML and AI. This critical decision-making data will be made available through modem cloud networking, access control, and cross-domain solutions to those who require access. Common data standards will be a key part of the Department’s methodology for tagging, storing, accessing, and processing information. Ensuring an enterprise cloud environment will increase the transparency of this data, and drive the velocity of data analysis, processing, and decision making. Leveraging advances in commercial cloud security technologies will ensure the Department’s information is protected at the appropriate level. (Shanahan, 2018, pp. 5–6)

Methodologies

Knowledge Value Added

Benefits

Knowledge Value Added (KVA) is a way for measuring the value produced by a system and its subprocesses that are objective and quantitative. Analysts can compare the obtained ratios to the ratios from other subprocesses to establish their relative efficacy because each process’ value measurements employ ratio scale numbers. KVA translates all process outputs into common value units, resulting in a consistent productivity performance ratio across all operations. PMs can compare the value added by IT processes to the value generated by the human component. PMs can use these measurements to build meaningful ratios in their study of the program’s performance thanks to the scales. Return on knowledge (ROK; i.e., a process’s common unit outputs) is divided by the process cost necessary to produce the outputs, and for ROI calculations, the ratio is monetized outputs minus cost divided by cost. The ROKs and ROIs, which are always 100% associated, inform managers about the amount of value a process provides versus the amount of money invested to achieve that value. Unlike any other methodology, KVA assigns these figures to both the process and subprocesses, not only the company as a whole (as is done in standard generally accepted accounting practice metrics used in standard financial ratios).

Conducting a KVA analysis of a program will provide a PM with a clearer understanding of the value of the program’s operational components. While most firms utilize cost/schedule metrics to assess the success of a project or operation, ROK will provide them with additional value-based data to help them make better management decisions. The relative predicted baseline value of the program’s components can be determined using PMs. Knowing that a certain job or subprocess produces the same output value as another process but at a different cost can help you understand why the entire system is performing differently. As a result, experienced managers have the information they need to dedicate resources to specific program components that need improvement or should be used more frequently, resulting in increased value-added. It also enables for estimations of the potential value-added of an AI system feature that was not originally planned for the project.

While a KVA study can provide information to aid in program or project management, it does not necessitate significant changes to organizational structure or reporting systems. Without bringing complicated new measures into the system, the review can be carried out as part of standard reporting procedures. The learning time, process instruction (e.g., WBS can be used as a surrogate for this technique), and binary query method are all dependent on data from the project description and requirements documents. To validate the accuracy of the presented data, a modest amount of hands-on measuring may be required. As a result, the analysis can be completed faster than other standard assessment approaches (e.g., activity-based costing), providing PMs with more timely access to relevant data.

Challenges

The value of the components that produce the outputs of the subprocesses will be quantified using KVA, which is a ratio-scale number. It does this, however, only with processes that have known a priori outputs. The intangible objects that occur within the human brain, such as creativity and imagination, cannot be quantified using this method, or any other method for that matter. In reality, because there is no formula for creativity, no present method can effectively quantify these types of intangibles within a process. Because the creative process cannot be learned or described algorithmically, these factors are not common to the ordinary user and, hence, cannot be specified using any of the KVA methods—learning time, binary query, or process description. Once creativity has been used to create an AI capacity, KVA can be used to algorithmically describe its productivity. KVA assigns a process's current value, but it can't forecast the value of potential future additional outputs unless they can be described using one of the KVA methods.

Although KVA will supply ratio-scale data to assist in analyzing processes inside a program, the ratios are frequently only useful for comparisons between projects. Benchmarking the raw figures against other organizations or other divisions within the same organization will give a useful benchmark for assessing predicted ROK performance. The resulting ROK and ROI measurements will be comparable among organizations (for business and non-profit) that create diverse products or services, regardless of the language used to describe outputs. Because these output descriptions are in standard units, they can be viewed as a value constant across all processes, with the value of a component subprocess or core process determined solely by the number of outputs. The end outcome of any correctly completed research will yield similar ROK and ROI estimations, which is KVA's ultimate purpose.

Integrated Risk Management

To forecast when various projects will be completed, all organizations rely largely on project planning software. Completing projects on schedule, on budget, and to a set value is crucial to the effective operation of a business. Many factors can influence a timetable in today's high-tech world. When it comes to technical capabilities, they frequently fall short of expectations. In many circumstances, requirements may be insufficient and require more elaboration. Tests might produce unexpected results, both good and harmful. Cost rises, timetable lapses, and value variations can all be caused by a variety of factors. In rare circumstances, we may be blessed with good fortune, and the schedule can be accelerated without jeopardizing the project's productivity.

Project timelines are inherently insecure, and changes are expected. As a result, we should anticipate changes and devise the best strategy for dealing with them. So why do projects take so much longer than expected? The inaccuracy of timetable estimation is one of the reasons. The following discussion describes the flaws in standard timetable estimation approaches, as well as how simulation and advanced analytics can be used to remedy these flaws.

It's crucial to first comprehend the Integrated Risk Management (IRM) process and how the various methodologies are related in the context of risk analysis and risk management. From a qualitative management screening process to provide clear and concise reports for management, this framework contains eight separate steps of a successful and complete risk analysis implementation. The author (Johnathan Mun) established the process based on past successful risk analysis, forecasting, real options, valuation, and optimization projects in both consultancy and industry-specific settings. These phases can be completed independently or in order for a more thorough integrated study.

The procedure can be broken down into eight easy steps (Mun, 2016):

- **Qualitative Management Screening**
- **Forecast Predictive Modeling**
- **Base Case Static Model**
- **Monte Carlo Risk Simulation**
- **Real Options Problem Framing**
- **Real Options Valuation and Modeling**
- **Portfolio and Resource Optimization**
- **Reporting, Presentation, and Update Analysis**

Qualitative Management Screening

The first stage in every IRM process is qualitative management screening. In accordance with the firm's mission, vision, goal, or overall business strategy, management must determine which projects, assets, initiatives, or strategies are viable for further analysis, which may include market penetration strategies, competitive advantage, technical, acquisition, growth, synergistic, or globalization issues. That is, the initial list of initiatives should be qualified in terms of how well they would achieve management's objectives. When management frames the entire problem to be solved, the most important insight is often generated. The numerous dangers to the firm are identified and flushed out in this step.

Forecast Predictive Modeling

If historical or comparable data is available, the future is projected using time-series analysis or multivariate regression analysis. Other qualitative forecasting methods may be employed instead (subjective guesses, growth rate assumptions, expert opinions, Delphi method, etc.). Future revenues, sale price, quantity sold, volume, production, and other key revenue and cost drivers are projected at this stage in the financial process. Time series, nonlinear extrapolation, stochastic process, ARIMA, multivariate regression forecasts, fuzzy logic, neural networks, econometric models, GARCH, and other methods are examples of methodologies.

Base Case Static Model

A discounted cash-flow model is generated for each project that passes the initial qualitative tests, whether it is for a single project or numerous projects under consideration (KVA analysis, using the market comparables approach, can be used to monetize value for this model). Using the anticipated values from the previous phase, a net present value is generated for each project using this model as the base case analysis. The traditional approach of modeling and forecasting revenues and expenses, then discounting the net of these revenues and costs at an appropriate risk-adjusted rate, yields this net present value. Here are calculated the return on investment, as well as other profitability, cost-benefit, and productivity indicators.

Monte Carlo Risk Simulation

Because the static discounted cash flow only provides a single-point estimate, there is often little trust in its accuracy, especially given the significant uncertainty surrounding future events that affect expected cash flows. Next, Monte Carlo risk simulation should be used to better evaluate the actual worth of a project. The discounted cash-flow model is normally subjected to a sensitivity analysis first; that is, by designating the net present value as the outcome variable, we can vary each of the previous variables and see how the resulting variable changes. As they go through the model, revenues, costs, tax rates, discount rates, capital expenditures, depreciation, and other prior factors all have an impact on the net present value number. By tracing back all of these previous variables, we can change each of them by a predetermined amount and assess the effect on the resulting net present value. Due to its shape, the most vulnerable preceding variables are depicted first, in descending order of magnitude, on a graphical depiction that is frequently referred to as a tornado chart. With this information, the analyst can evaluate which crucial aspects are deterministic in the future and which are very uncertain. The uncertain important variables that drive the net present value and, thus, the decision are known as critical success drivers. For these critical success criteria, Monte Carlo simulation is an excellent fit. Because several of these critical success determinants are linked—for example, operational costs may rise in proportion to the quantity sold of a particular product, or prices may be inversely associated to quantity sold—a correlated Monte Carlo simulation may be required. The majority of the time, historical data can be used to make these relationships. When you run correlated simulations, you get a lot closer to the real-world behavior of the variables.

Real Options Problem Framing

The dilemma now is what to do after quantifying hazards in the previous stage. The risk data gathered must be transformed into actionable intelligence in some way. So what, and what do we do about it, just because risk has been estimated as such and such using Monte Carlo simulation? The solution is to apply actual options analysis to mitigate these risks, value them, and position yourself to profit from them. The act of defining the problem generates a strategic map, which is the first stage in real possibilities. Certain strategic options for each project would have been obvious based on the overall problem identification that occurred during the initial qualitative management screening phase. The strategic options could include, for example, the ability to expand, contract, abandon, switch, choose, and so on. The analyst can then choose from a list of choices to investigate further based on the identification of strategic options that exist for each project or at each stage of the project. Real options are incorporated into projects to protect against downside risks and to profit from upswings.

Real Options Valuation and Modeling

The resulting stochastic discounted cash-flow model will have a distribution of values thanks to Monte Carlo risk simulation. As a result, simulation models, analyzes, and quantifies each project's unique risks and uncertainties. As a result, the NPVs and project volatility are distributed. We assume that the underlying variable in real options is the project's future profitability, which is represented by the future cash-flow series. The results of a Monte Carlo simulation can be used to calculate the implied volatility of the future free cash flow or underlying variable. Usually, the volatility is measured as the standard deviation of the logarithmic returns on the free-cash-flow stream (other approaches include running GARCH models and using simulated coefficients of variation as proxies). Furthermore, in real options modeling, the present value of future cash flows for the base case discounted cash-flow model is used as the initial underlying asset value. Real options analysis is used to determine the strategic option values for the projects using these inputs.

Portfolio and Resource Optimization

Portfolio optimization is a step in the analysis that can be skipped. Because the projects are usually associated with one another, management should view the results as a portfolio of rolled-up projects if the analysis is done on numerous projects. Viewing them individually will not offer the actual picture. Because businesses don't just have one or two initiatives, portfolio optimization is essential. Because certain projects are interconnected, there is potential for risk hedging and diversification through a portfolio. Portfolio optimization takes all of these factors into account to build an optimal portfolio mix because firms have limited budgets, as well as time and resource constraints, while also having needs for particular overall levels of returns, risk tolerances, and so on. The research will determine the best way to allocate funds across multiple projects.

Reporting, Presentation, and Update Analysis

Until reports can be created, the analysis is not complete. Not only should the results be communicated, but so should the process. A complex black box set of analytics is transformed into transparent processes by clear, simple, and exact explanations. Management will never accept outcomes from black boxes if they don't know where the assumptions or data come from or what kind of mathematical or financial manipulation is going on. Risk analysis presupposes that the future is uncertain and that management has the authority to make mid-course corrections when these uncertainties or risks are resolved; the analysis is typically performed ahead of time and, therefore, ahead of such uncertainty and risks. As a result, if these risks are identified, the analysis should be updated to integrate the decisions made or to revise any input assumptions. Several iterations of the real options analysis should be undertaken for long-horizon projects, with future iterations being updated with the newest data and assumptions.

Understanding the processes required to complete the IRM process is critical because it reveals not only the technique itself but also how it differs from previous analyses, indicating where the traditional approach finishes and the new analytics begin.

Benefits

IRM is a great tool for improving the quality of information accessible while making decisions because it combines multiple proven strategies. When applied to the examination of potential initiatives and investments, dynamic Monte Carlo simulation depicts the risks connected with the projects in a more realistic manner than traditional methodologies. Static forecasting based on assumptions and past performance provides a restricted view of a project's potential outcomes. Decision-makers can acquire a more full understanding of the project's uncertainty by running thousands of simulations or more while altering the variables within realistic possibilities. Increasing the amount of relevant and correct information available to managers will increase the quality of the leadership team's decisions.

IRM takes a methodical strategy to deal with AI investments. Following the eight phases is a simple procedure that aids in the quantitative decision-making process. While the functions within each phase can be sophisticated and require additional training, the overall process is straightforward and simple to follow. Because the IRM approach is fully defined, it may be integrated into existing procedures without requiring a complete reengineering. IRM will use data from existing approaches and expand it to improve the scope of a project's evaluation. The true possibilities were quantified, and the outcome diverged from what was expected. The systemic design of IRM allows different members or teams to finish the process without having to re-collect data and start from the beginning. Analysts should be able to continue the procedure from any point in the approach after completing IRM training.

Real options analysis provides managers with the probability of certain project results, allowing them to select the best way to proceed with a project. Real options were offered not only at the start of the program, with three different routes in which the program may go, but also at each stage of the chosen strategy. By drafting a contract that allows an organization to modify its course of action as more information becomes available, the corporation can reduce losses from failing programs while maximizing gains from initiatives that are succeeding or showing promise. Fortunately, many viable possibilities are already ubiquitous in DoD buys. Contracts are frequently canceled by the government due to changes in budgetary policy, inability to satisfy requirements, or other factors. Including other genuine choices in contracts isn't an entirely new concept.

The use of common units to make strategic decisions about a system's value is a core component of the IRM methodology. Leadership can see a statistical range reflecting the potential value of a project by incorporating KVA values into the static and dynamic IRM models. The present values of the genuine option strategies were calculated using the market comparable prices produced by the value analysis. The effectiveness of most other ways is determined only by the program's cost, presuming that the value is inherent owing to the needs that were produced. IRM can provide decision-makers with information on both the expenses of a proposed investment in an initiative and the value of that project in comparable units.

Challenges

While IRM is a very useful analytical tool, it does have some disadvantages. The method's multiple techniques might be challenging to master (Housel et al., 2019). To do a full study, it is a hard process that necessitates a solid understanding of both finance and statistics. While computing tools can help with the analysis, the inputs are more involved than simply typing a few numbers into a program and receiving the results. An analyst can generate the essential information to enable decision-makers access to the proper comparison material to make an informed decision if they have a good understanding of the core principles, enough training, and the right tools (Housel et al., 2019). The amount of data gathered during statistical analysis can be overwhelming. The simulations and their conclusions appear to originate from a quantitative black box to individuals without a strong statistics background (Mun, 2016). If decision-makers don't comprehend why an analyst makes a recommendation, it's simple to dismiss the advice and fall back on tried-and-true methods. To tackle this possible issue, create detailed and complete reports for management review, as well as knowledgeable presentations to allay worries about the unfamiliar procedures. To take advantage of actual options, they must be reviewed before a decision is made to implement any of them. When writing contracts, leadership must consider the future option to ensure that certain alternatives stay available. Some alternatives, such as expanding, can be implemented very easily by building a new project based on the first investment's success. However, if the contract does not include relevant conditions, project managers may not have as much flexibility in abandoning the project. Vendors must be willing to accept the possibility of subcontract cancellation when they are not at fault, which may increase the cost of completing a task. Managers must perform a careful study of which prospective options may be exercised in the future before signing contracts with vendors, due to the potential increased cost associated with contracting genuine options.

IRM, like all financial forecasting, makes projections based on previous data. Decision-makers can gain more insight from predictions that incorporate current information rather than relying just on historical trends. Meteorologists, for example, compile weather forecasts from a variety of sources: Current weather conditions are monitored using Doppler radar, satellites, radiosondes (weather balloons), and automated surface-observing systems (National Oceanic and Atmospheric Administration [NOAA], 2017). The data from multiple sources is run in models based on known historical patterns for the region using numerical weather prediction (NOAA, n.d.). Knowing the present conditions is just as crucial to a meteorologist as knowing the past models (NOAA, n.d.). Similarly, the models would deliver even more precise information if the project analyst could add pertinent information that is up to the minute (or to the requisite quality). Because of previous projects with historical data, outsourcing, lowering manning and retaining the current structure all offer statistics that could be used in simulations. Despite the fact that this weakness is not exclusive to the IRM technique, executives should be aware of it in any financial forecast.

Finally, the DoD not currently reward PMs who reap the rewards of risk. The risk framework in DoD acquisitions is intended to reduce project costs and schedule overruns. DoD contracts are structured in such a way that they do not incentivize vendors or the project as a whole to improve their capabilities or performance. When a for-profit company invests in an initiative that may fail, it does so because the potential upside gain outweighs the risk of failure. For example, if an aircraft's design target is to attain 250 knots and the design threshold is 200 knots, the budget will be allocated to the threshold rather than the objective. Unless the PM is able to reallocate resources internally, the program will not be able to meet its objectives. The acquisitions process considers the cost of achieving the goal rather than the worth of the goal. Performance is rewarded in for-profit businesses, which is evaluated by revenue. The DoD's implicit surrogate for revenue is cost reductions, which has a different value than improving a project's worth. Acquisitions by the DoD, on the other hand, are only made when the negative implications have been mitigated to the maximum extent practicable. The upside risk is unimportant to the PMs; all that matters is that the program is finished on time and on budget. Although it is still important to look at how potential projects fit into the DoD's broader collection of acquisitions and current assets, the contract structure limits certain IRM portfolio optimization features.

Comparison of Key Attributes

The type of methodology to use should be determined by the nature of the project at hand, including the level of commitment required from the organization, the organization's desire to align strategic goals with the project, the methodology's predictive capability, the flexibility required, and the amount of time available. While others in the business must understand concepts in order to comprehend status reports, EVM just requires the management team to track the project's cost and schedule against the baseline because there is no pre-determined goal alignment with the organization. While the CPI and SPI can assist in estimating the ultimate cost and schedule, EVM has no true predictive potential because it is assumed that the schedule would follow the baseline regardless of historical performance volatility. In EVM, sticking to the baseline is critical, and altering requirements can substantially affect the baseline, lowering the methodology's effectiveness. For an AI project with its many unknown components and capabilities a priori, setting up, monitoring, and reporting the cost/schedule performance of each work item inside the WBS can be a time-consuming and costly operation.

To assess the value of a process or component output, KVA simply requires the KVA analyst and the process owner, who serves as the SME, supporting the requirement to match the project with an organization's productivity goals. They can model the present baseline as-is process ROK and compare it to the proposed to-be process model ROK using this approach, resulting in a straightforward forecast of the improvement between the models. Because KVA can be used with any description language that defines process outputs in common units, analysts can choose the method that is most helpful for the system in question, allowing for flexibility. This analysis may be conducted fast, with a rough assessment available in a few days. To assess how a project fits into an organization's portfolio, the project's present value (PV), and potential real possibilities, IRM requires organizational leadership, portfolio and project managers, and the analyst. IRM gives a prediction of a project's anticipated performance by analyzing and simulating alternative situations, allowing managers to build in flexibility via genuine options at the right spots within the project. Assuming that the data required for the analysis is available, the process can be done quickly.

Methodologies in AI Acquisition

As previously stated, each methodology has strengths and weaknesses that make it more appropriate for certain applications than others. The iterative nature of software development is the most difficult aspect of adopting EVM when gaining AI. To be most successful, EVM requires well-stated, specific requirements for intermediate phases. While software program outputs are well specified, the methods required to produce the software are not, causing challenges when estimating cost and schedule. EVM can adequately monitor the progress if the software is not complex or comprises well-known operations. Integrating software and hardware is also difficult with EVM since there are various elements of the program that must be merged to achieve the objectives, requiring additional debugging and recoding. When used to manage the physical production of systems or infrastructure, EVM is more efficient. It can track the cost and schedule progress of software work packages, but it's not as good at determining their worth.

Any IS system can use KVA to offer an objective, ratio-scale measure of value and cost for each core process and its subprocesses or components. Managers can then examine productivity ratios information, such as ROK and ROI, using the two factors to determine the efficiency of a process in relation to the resources utilized to create the output. This can assist the manager in deciding how to allocate resources for system updates or estimating the future value of a system that is being purchased. Managers can iterate the value of system real options analysis using simulation and other ways by combining KVA and IRM data. IRM can also use past data to evaluate risks and anticipate performance probability for metrics of potential success for programs and program components. It's a tool that can help with investment strategy and can be used to acquire any form of AI. It is not, however, intended to assist in the procurement of an AI program or in determining how to meet the program's specific criteria.

Summary

The scope, capabilities, and limitations of various AI systems are demonstrated by examining the benefits and challenges of the proposed approaches. It also aids in determining which areas and phases of the Defense Acquisition System life cycle the methodologies or components of the methodologies should be included. The following section offers suggestions based on the findings.

Conclusion

Simply put, how might certain advanced analytical decision-making processes be applied in the acquisition life cycle to supplement existing procedures to ensure a successful acquisition of AI technologies?

As previously stated, EVM is the sole program management methodology that the U.S. government requires for all DoD acquisition initiatives worth more than $20 million. Regardless of this necessity, EVM is a methodology that offers a systematic approach to IT acquisition through program management processes that can assist in keeping an acquisition program on track and below estimated cost estimates. However, there are substantial drawbacks to utilizing EVM for AI acquisitions, the most prominent of which is that it was not built to manage AI acquisitions that follow a highly iterative and volatile course. Organic AI acquisitions necessitate a high level of flexibility in order to deal with the unknowns that surface during the development process, as well as value-adding opportunities that were not anticipated. Furthermore, EVM lacks a uniform unit of value metric that would allow typical productivity metrics like ROI to be calculated. When a program's worth is determined by how closely it adheres to its initial cost and schedule projections, the program's performance may suffer in terms of output quality when intended program activities become iterative, as in the development of many AI algorithms. EVM is not designed to recognize disproportionate increases in value if an AI acquisition program is going toward cost and schedule overruns, but the ensuing value-added of the modifications to the original requirements offers disproportionate increases in value.

To address EVM's shortcomings in AI acquisitions, the methodology should be combined with KVA and IRM, which can be useful during the EVM requirements and monitoring phases by ensuring that a given AI acquisition is aligned with organizational strategy and that a baseline process model has been developed for establishing current performance prior to the acquisition of an AI system. After the AI has been obtained, a future process model that forecasts the value-added of incorporating the AI can be used to set expectations that can be tested against the baseline model. IRM can be used to anticipate the value of strategic real choices flexibility that an acquired AI might bring, allowing leadership to choose the alternatives that best meet their desired goals for AI in defense core activities.

KVA should be utilized in AI acquisitions because it gives an objective, quantitative measure of value in common units, allowing decision-makers to better comprehend and compare different strategic options based on their value and cost. Only by employing KVA to determine the value inherent in the system can AI systems be given a return on investment. PMs benefit from this information since it gives them a more full picture of the current and future systems' performance.

When obtaining AI through the Defense Acquisition System, it's also a good idea to use IRM. The risk estimates associated with the components and subcomponents of a program, in terms of potential cost overruns, value variabilities, and schedule delays, can be improved by using dynamic and stochastic uncertainty and risk-based modeling techniques to predict likely and probabilistic outcomes. Analyzing multiple real-world options in the context of the models' outputs will assist PMs in making the best decisions possible when defining the future of a program.

As is now done, PMs should only employ EVM throughout the Engineering and Manufacturing Development (EMD) phase. EVM, on the other hand, will operate best in hardware manufacturing solutions with fully mature technology prior to the program's start. EVM is not well suited for AI development because many AI acquisition efforts involve upgrading current technology and generating new software solutions to meet requirements. Nonetheless, PMs can employ a variety of agile EVM strategies to complete projects on time and on budget if the proper procedures are done when establishing the baseline. Requirements must be broken down into tiny, simply defined tasks, with risk and uncertainty elements appropriately accounted for in the timetable. Other approaches, such as KVA and IRM, should be used in conjunction with EVM to guarantee that these elements are based on verifiable measurements rather than assuming how much more time, money, and value may be required to execute complex tasks.

KVA and IRM will assist in determining the value of the various options evaluated in the analysis of alternatives (AoA) process during the Materiel Solution Analysis (MSA) phase. KVA can objectively assess the value of the current, as-is system as well as potential future systems. Then IRM can leverage other aspects like cost, value, complexity, and schedule to value the alternatives in terms of their respective parameter values. As the chosen solutions mature during the TMRR phase, a revised KVA analysis will reassess initial estimations and provide a predicted ROI that may be incorporated into an IRM risk and actual alternatives analysis for the AI solution before entering the EMD phase, if necessary.

Limitations and Future Research

This study looked into whether the various methodologies—EVM, KVA, and IRM—could be used to improve AI acquisition inside the Defense Acquisition System. Future research should look at how these approaches interact with or improve other acquisition system components. This comprises the specific procedures of JCIDS and PPBE, as well as the interactions between JCIDS, PPBE, and the Defense Acquisition System as a whole. Certain approaches, such as IRM, may be more useful when applied to the full acquisition process rather than just a part of it. Future research might also look into how these diverse methods could be utilized to acquire things that aren't related to AI or IT.

The study focused on AI as a whole, rather than individual types of AI. Future research should look into whether acquisition methods, strategies, and methodologies differ depending on the type of AI being acquired. This is particularly relevant when it comes to artificial intelligence and its subsets. Based on their complexity, intricate nature, developing technology, and amount of risk, machine learning, intelligence with a specific emphasis or field of specialty, and general or universal intelligence will likely use different ways in the acquisition process.

Another area of prospective investigation is the use of these approaches in commercial AI acquisition. The focus of this study was solely on the application of the strategies in the DoD acquisition process. Commercial entities, on the other hand, face challenges when adopting extensive or complicated AI and IT systems, especially when the technologies are used at the enterprise level. Further research may reveal whether these same techniques could be useful to private-sector decision-makers during the development, adoption, or customization of commercial AI. The hype cycle for AI and automation is on the rise, as highlighted in the literature, and the demand to buy such technology is as relevant for the private sector as it is for the DoD. In addition, the current pandemic triggered by Coronavirus Disease 2019 (COVID-19) has compelled a permanent shift in society toward permanent distant labor. Because these trends are expected to continue in the near future, more automation tools will be needed to support this workforce. As part of the Fourth Industrial Revolution and Industry 4.0, these developments could be investigated for their consequences.

Finally, this study looked at only the most promising approaches out of a wide range of options. Other program management tools, management philosophies, analytic tools, or other approaches, as well as their benefits while adopting AI, should be investigated in future research. While the approaches investigated were chosen because they are likely to enhance the process and assist EVM improvements, other systems may be more appropriate in certain phases or provide additional benefits not seen in this study.

References

Betts, K. D., & Jaep, K. R. (2017). The dawn of fully automated contract drafting: Machine learning breathes new life into a decades-old promises. *Duke Law and Technology Review*, *15*(1), 216–233.

Darken, R. (2019, October 21). *Human–machine teaming AI*. Naval Postgraduate School.

DoD. (n.d.). *PALT*. https://DoDprocurementtoolbox.com/site-pages/palt

DoD. (2003, May 12). *The defense acquisition system* (change 2, DoD Directive 5000.01). https://www.esd.whs.mil/Directives/issuances/DoDd/

DoD. (2005). *Work breakdown structures for defense materiel items* (MIL-HDBK-881A). http://www.acqnotes.com/Attachments/MIL-Hand%20Book-881.pdf

DoD. (2011, April). *Technology Readiness Assessment (TRA) guidance*. Assistant Secretary of Defense for Research and Engineering.

DoD. (2014, March 12). *Risk management framework (RMF) for DoD information technology (IT)* (change 2, DoD Instruction 8510.01). https://www.esd.whs.mil/Directives/issuances/DoDi/

DoD. (2015a, January 7). *Operation of the defense acquisition system* (change 4, DoD Instruction 5000.02). https://www.esd.whs.mil/Directives/issuances/DoDi/

DoD. (2015b, September). *DoD program manager's guide for integrating the cybersecurity risk management framework (RMF) into the system acquisition lifecycle*.

DoD. (2017, February 26). *Defense acquisition guidebook*. https://www.dau.mil/tools/dag

DoD. (2018, April 16). *Agile and earned value management: A program manager's desk guide*.

DoD. (2019a, January 18). *Earned value management implementation guide (EVMIG)*.

DoD. (2019b, February 12). *Summary of the 2018 Department of Defense artificial intelligence strategy: Harnessing AI to advance our security and prosperity*. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Eykholt, K., Evitimov, I., Fernandes, E., Li, B., Rahmati, A., Xia, C., & Song, D. (2018). *Robust physical-world attacks on deep learning visual classification*. Cornell University.

Gardner, H. (1993). *Multiple intelligences*. Basic Books.

Golstein, B. (2018, October 10). *A brief taxonomy of AI*. SharperAI. https://www.sharper.ai/taxonomy-ai/

Greenfield, D. (2019, June 19). *Artificial intelligence in medicine: Applications, implications, and limitations*. Harvard University. http://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/

Gunning, D. (2017, November). *Explainable artificial intelligence*. DARPA. https://www.darpa.mil/attachments/xaiprogramupdate.pdf

Housel, T., & Bell, A. H. (2001). *Measuring and managing knowledge*. McGraw Hill.

Housel, T., & Kanevsky, V. A. (1995). Reengineering business processes: A complexity theory approach to value added. *Infor*, *33*(4), 248. https://doi.org/10.1080/03155986.1995.11732285

Housel, T., & Kanevsky, V. (2006). *Measuring the value added of management: A knowledge value added approach*. Naval Postgraduate School.

Housel, T. J., Mun, J., Jones, R., & Carlton, B. (2019a, May). *A comparative analysis of advanced methodologies to improve the acquisition of information technology in the Department of Defense for optimal risk mitigation and decision support systems to avoid cost and schedule overruns* [Paper presentation]. Sixteenth Annual Acquisition Research Symposium, Monterey, CA, United States.

Housel, T. J., Mun J., Jones, R., & Carlton, B. (2019b, October). *A comparative analysis of advanced methodologies to improve the acquisition of information technology in the Department of Defense for optimal risk mitigation and decision support systems to avoid cost and schedule overruns* (NPS-AM-20-002). Naval Postgraduate School. https://dair.nps.edu/handle/123456789/2774

Icertis. (n.d.-b). *Customers*. Retrieved October 11, 2019, from https://www.icertis.com/customers/

Icertis. (2019a). *Icertis customer profile: Mindtree*. https://www.icertis.com/customer/mindtree/

Icertis. (2019b, April 2). *Microsoft streamlined its contract management*. https://www.icertis.com/customers/microsoft-information-exchange-agreements-case-study/

Institute for Robotic Process Automation & Artificial Intelligence. (2019, November 16). *What is robotic process automation?* https://irpaai.com/what-is-robotic-process-automation/

King, A. D. (2019, November 16). *Talk to transformer*. https://talktotransformer.com/

Knight, W. (2017, April 11). *The dark secret at the heart of AI*. Technology Review. https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

Lawgeex. (2018, February). *Comparing the performance of artificial intelligence to human lawyers in the review of standard business contracts*. Law.com. https://images.law.com/contrib/content/uploads/documents/397/5408/lawgeex.pdf

Leviathan, Y. (2018, May 8). *Google Duplex: An AI system for accomplishing real-world tasks over the phone*. Google AI Blog. https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html

Mun, J. (2015). *Readings in certified quantitative risk management (CQRM)* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2016a). *Real options analysis* (3rd ed.). Thomson-Shore and ROV Press.

Mun, J. (2016b, October). *Empirical cost estimation tool* [Paper presentation]. Naval Acquisitions Research Conference, Monterey, CA, United States.

Mun, J. (2019). Empirical cost estimation for U.S. Navy ships. *Universal Journal of Management*, *7*, 152–176.

Mun, J., & Anderson, M. (2021, January). Technology trust: System information impact on autonomous systems adoption in high-risk applications. *Defense Acquisition Research Journal, 28*(1), 2–39. https:// doi.org/10.22594/10.22594/dau.19-841.28.01

Mun, J., George, K., & Ledbetter, E. (2020). *Total ownership with life-cycle cost model under uncertainty for surface ships' electro-optical-infrared-sensors* [Unpublished manuscript].

Mun, J., & Housel, T. (2010). *A primer on applying Monte Carlo simulation, real options analysis, knowledge value added, forecasting, and portfolio optimization.* Calhoun.

Mun, J., Housel, T., & Wessman, M. D. (2010). PEO-IWS ACB insertion portfolio optimization. In *Proceedings of the Seventh Annual Acquisition Research Symposium* (Vol. 2, No. NPS-AM-10-069-VOL-2). Naval Postgraduate School. https://my.nps.edu/documents/105938399/108624025/NPS-AM-10-069.pdf/c71c6830-853a-448b-beac-242bea4cba8b?version=1.0

Nayak, P. (2019, October 25). *Understanding searches better than ever before*. Google. https://blog.google/products/search/search-language-understanding-bert

Parloff, R. (2016, September 28). From 2016: Why deep learning is suddenly changing your life. *Fortune.* https://fortune.com/longform/ai-artificial-intelligence-deep-machine-learning/

Raghaven, S., & Mooney, R. J. (2013). *Online inference-rule learning from natural-language extractions*. The University of Texas.

Rich, B. (2018, February 12). How AI is changing contracts. *Harvard Business Review.* https://hbr.org/2018/02/how-ai-is-changing-contracts

Sammalkorpi, S., & Teppala, J. P. (2019). *AI in procurement*. Sievo Oy.

Shanahan, P. (2018). *DoD cloud strategy*. Department of Defense.

Sievo. (2019, November 16). *AI in procurement*. https://sievo.com/resources/ai-in-procurement

Zarkadakis, G. (2019, September 11). *The rise of the conscious machines: How far should we take AI?* Science Focus. https://www.sciencefocus.com/future-technology/the-rise-of-the-conscious-machines-how-far-should-we-take-ai/

Acquisition Research Program
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net