NPS-PM-21-014

ACQUISITION RESEARCH PROGRAM
SPONSORED REPORT SERIES

Acquiring Artificial Intelligence Systems: Development Challenges, Implementation Risks, and Cost/Benefits Opportunities

October 27, 2020

Dr. Johnathan Mun, Professor of Research

Dr. Thomas Housel, Professor

COL Raymond Jones (Ret.), Professor of Practice

Dr. Timothy Shives, Adjunct Faculty

Major Benjamin Carlton, USMC, Student

Major Vladislav Skots, USA, Student

Graduate School of Operational and Information Sciences

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

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

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact the Acquisition Research Program (ARP) via email, arp@nps.edu or at 831-656-3793.

Abstract

The acquisition of artificial intelligence (AI) systems is a relatively new challenge for the U.S. Department of Defense (DoD). Given the potential for high-risk failures of AI system acquisitions, it is critical for the acquisition community to examine new analytical and decision-making approaches to managing the acquisition of these systems in addition to the existing approaches (i.e., Earned Value Management, or EVM). In addition, many of these systems reside in small start-up or relatively immature system development companies, further clouding the acquisition process due to their unique business processes when compared to the large defense contractors. This can lead to limited access to data, information, and processes that are required in the standard DoD acquisition approach (i.e., the 5000 series). The well-known recurring problems in acquiring information technology automation within the DoD will likely be exacerbated in acquiring complex and risky AI systems. Therefore, more robust, agile, and analytically driven acquisition methodologies will be required to help avoid costly disasters in acquiring these kinds of systems. This research provides a set of analytical tools for acquiring organically developed AI systems through a comparison and contrast of the proposed methodologies that will demonstrate when and how each method can be applied to improve the acquisitions life cycle for AI systems, as well as provide additional insights and examples of how some of these methods can be applied. This research identifies, reviews, and proposes advanced quantitative, analytically based methods within the integrated risk management (IRM) and knowledge value added (KVA) methodologies to complement the current EVM approach. This research examines whether the various methodologies—EVM, KVA, and IRM—could be used within the Defense Acquisition System (DAS) to improve the acquisition of AI. While this paper does not recommend one of these methodologies over the other, certain methodologies, specifically IRM, may be more beneficial when used throughout the entire acquisition process instead of within a portion of the system. Due to this complexity of AI system, this research looks at AI as a whole and not specific types of AI.

THIS PAGE LEFT INTENTIONALLY BLANK

About the Authors

Dr. Tom Housel specializes in valuing intellectual capital, knowledge management, acquisitions research, telecommunications, information technology, value-based business process reengineering, and knowledge value measurement in profit and nonprofit organizations. He is currently a tenured full professor for the Information Sciences (Systems) Department of the U.S. Naval Postgraduate School (NPS). He has managed a $5+ million portfolio of field studies, educational initiatives, and industry relationships.

His current research focuses on determining the most promising methodologies for improving the acquisitions life cycle and developing a physics-based model to predict revenue growth in knowledge-intensive companies. He continues to work in the behavioral accounting decision-making area and has published in a number of international accounting journals (e.g., ABR, EAR). His research was published twice in the top information systems journal, *MIS Quarterly* (MISQ).

Prior to joining NPS, he was an associate professor at the Marshall School of Business at the University of Southern California. His last assignment in the corporate world was as the chief of consumer market research for Telecom Italia in Venice, Italy, where he developed new methods for predicting the adoption rates for new interactive multimedia broadband applications.

He received his PhD from the University of Utah in 1980. He won the prestigious Society for Information Management award for best paper in the field in 1986 that was subsequently published in the MISQ. His work on measuring the value of intellectual capital has been featured in a *Fortune* cover story (October 3, 1994) and in numerous books, professional periodicals, and academic journals. His books include *Measuring and Managing Knowledge* (English and Chinese versions) and *Global Telecommunications Revolution: The Business Perspective* with McGraw-Hill (both in 2001).

Dr. Johnathan C. Mun is a research professor at NPS (Monterey, CA) and teaches master's and doctoral courses as well as executive seminars in quantitative risk analysis, decision sciences, real options, simulation, portfolio optimization, and other related concepts. He has also researched and consulted on many Department of Defense and Department of Navy projects and is considered a leading world expert on risk analysis and real options analysis. He has over 23 patents and patents pending and has authored 22 books including *Modeling Risk: Applying Monte Carlo Simulation, Real Options, Optimization, and Forecasting*, First, Second, and Third Editions (Wiley 2006, Wiley 2010, and Thomson-Shore 2015); *Real Options Analysis: Tools and Techniques*, First, Second, and Third Editions (Wiley 2003, Wiley 2005, and Thomson-Shore 2016); *Advanced Analytical Models* (Wiley 2008; ROV Press 2016); *The Banker's Handbook on Credit and Market Risk* (Elsevier Science 2008); and others. He is the creator of the following software: Real Options Super Lattice Solver, Risk Simulator, Project Economics Analysis Tool (PEAT), Modeling Toolkit, Risk Explorer, and ESO Valuation. His books and software are being used at top universities around the world (including the Bern Institute in Germany, Chung-Ang University in South Korea, Georgetown University, ITESM in Mexico, Massachusetts Institute of Technology, NPS, New York University, Stockholm University in Sweden, University of the Andes in Chile, University of Chile, University of Pennsylvania Wharton School, University of Hull in the United Kingdom, and Edinburgh University in Scotland).

Dr. Mun has taught at universities all over the world, from NPS and University of Applied Sciences (Switzerland and Germany) as full professor, to Golden Gate University (California) and St. Mary's College (California), and has chaired many graduate research thesis committees. He also teaches risk analysis, real options analysis, and risk analysis for managers public courses where participants can obtain the Certified in Quantitative Risk Management (CQRM) designation on completion of the week-long program. He also holds the position of the EU President of the American Academy of Financial Management and sits on the Global Board of Standards at the AAFM. He is currently the CEO of Real Options Valuation, Inc., and was formerly the vice president of analytics at Crystal Ball/Decisioneering, Inc., where he headed the development of options and financial analytics software products, analytical consulting, training, and technical support. Prior to that, he was a consulting manager and financial economist in the Valuation Services and Global Financial Services practice of KPMG Consulting and a manager with the Economic Consulting Services practice at KPMG LLP. He has extensive experience in econometric modeling, financial analysis, real options, economic analysis, and statistics. He has consulted on a variety of real options, risk analysis, financial forecasting, project management, and financial valuation issues for over 100 multinational firms (former clients include 3M, Airbus, Boeing, BP, Chevron Texaco, Financial Accounting Standards Board, Fujitsu, GE, Microsoft, Motorola, Pfizer, Timken, U.S. Department of Defense, U.S. Navy, Veritas, and many others). His experience prior to joining KPMG included being department head of financial planning and analysis at Viking Inc. of FedEx, performing financial forecasting, economic analysis, and research.

Dr. Mun received his PhD in Finance and Economics from Lehigh University, where his research and academic interests were in the areas of Investment Finance, Econometric Modeling, Financial Options, Corporate Finance, and Microeconomic Theory. He also has an MBA in business administration, an MS in management science, and a BS in Biology and Physics. He is Certified in Financial Risk Management (FRM), Certified in Financial Consulting (CFC), and Certified in Quantitative Risk Management (CQRM). He is a member of the American Mensa, Phi Beta Kappa Honor Society, and Golden Key Honor Society as well as several other professional organizations, including the Eastern and Southern Finance Associations, American Economic Association, and Global Association of Risk Professionals. Finally, he has written many academic articles published in the *Advances in Quantitative Finance and Accounting*, the *Global Finance Journal*, the *International Financial Review*, the *Journal of Financial Analysis*, the *Journal of Applied Financial Economics*, the *Journal of International Financial Markets, Institutions and Money*, the *Financial Engineering News*, and the *Journal of the Society of Petroleum Engineers*.

Professor Ray Jones retired as a colonel from the U.S. Army in 2012 and is a professor of practice with the Graduate School for Business and Public Policy at NPS. His last assignment in the Army was as the deputy program executive officer for the Joint Tactical Radio System (JTRS). Additionally, he served as the military deputy for the director of acquisition resources and analysis in the Office of the Under Secretary of Defense for Acquisition Technology and Logistics (USD[AT&L]), and managed three

Major Defense programs for the DoD, in addition to his many operational and research and development assignments. He graduated from the U.S. Naval Test Pilot School in 1995 and is 1983 graduate of the United States Military Academy. He has a Bachelor of Science degree in Aerospace Engineering, a Master of Science degree in Aeronautical Engineering from NPS, a Master's in Business Administration from Regis University, and a Master's Degree in National Resource Strategy from the Industrial College of the Armed Forces, and is currently a PhD candidate with the Graduate School of Information Sciences at NPS.

Dr. Timothy Shives is a Department of Defense (DoD) civil servant in the field of Information Technology (IT) management serving at the Defense Manpower Data Center (DMDC). Dr. Shives is also an officer in the United States Marine Corps and has served for more than 15 years in both the active duty and reserve components. Dr. Shives started his military career as an attack helicopter pilot flying the AH-1W "Super Cobra," but was then assigned to a multitude of other occupational specialties in the fields of financial management, information technology management, project management, acquisition management, and data management. After transitioning from active duty, Dr. Shives entered the DoD civil service as an information technology manager responsible for the program management and oversight of numerous large-scale, DoD-wide IT systems. Dr. Shives is currently on a special assignment to NPS's Information Sciences Department on an experiential detail from his parent organization as part of the Defense Senior Leadership Development Program (DSLDP), the DoD's premier leadership program for developing future senior executives (SES) from high-performing managers. The DSLDP candidates are selected from a pool of civil service managers throughout the department in a highly competitive selection process.

Dr. Shives holds a Doctor of Education (EdD) in Educational Leadership and an Education Specialist (EdS) postgraduate degree in Educational Leadership from Liberty University, a Master of Science (MS) in Information Technology Management (ITM) and a Master of Business Administration (MBA) from NPS, and a Master of Arts (MA) in National Security and Strategic Studies from the U.S. Naval War College. Dr. Shives also holds or has held the following professional certifications: Project Management

Professional (PMP), Certified Information Systems Security Professional (CISSP), Defense Acquisition Workforce Improvement Act (DAWIA) Program Management Level II, Certified Defense Financial Manager with Acquisition (CDFM-A), and Lean Six Sigma Black Belt (LSSBB). Finally, Dr. Shives just began his studies for his second doctorate, a Doctor of Philosophy (Ph.D.) in Information Sciences at NPS.

THIS PAGE LEFT INTENTIONALLY BLANK

NPS-PM-21-014

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Acquiring Artificial Intelligence Systems: Development Challenges, Implementation Risks, and Cost/Benefits Opportunities

October 27, 2020

Dr. Johnathan Mun, Professor of Research

Dr. Thomas Housel, Professor

COL Raymond Jones (Ret.), Professor of Practice

Dr. Timothy Shives, Adjunct Faculty

Major Benjamin Carlton, USMC, Student

Major Vladislav Skots, USA, Student

Graduate School of Operational and Information Sciences

Naval Postgraduate School

THIS PAGE LEFT INTENTIONALLY BLANK

Executive Summary

This report seeks to address the emerging field that falls under the umbrella term *artificial intelligence* (AI), and the Defense Acquisition System (DAS) that is not designed to develop and procure such state-of-the-art, rapidly evolving AI technologies. Thus, the research examines the challenges of acquiring an AI-based system within the typical acquisitions framework in business and in the Department of Defense (DoD) by conducting an analysis on a set of recommended quantitative tools for use in analyzing the processes in the acquisition of organically developed DoD AI systems.

The first section, "Introduction," discusses the research's problem statement, research objective, and research question. After a brief discussion on risk, the proposed three technical approaches evaluating DoD AI systems, Integrated Risk Management (IRM), Knowledge Value Added (KVA), and Earned Value Management (EVM) are introduced. The second section, "Literature Survey," provides a brief synthesis of the numerous articles surveyed, with an emphasis on the Defense Acquisition System, followed by the various approaches: EVM, KVA, and IRM. The third section, "Artificial Intelligence," delves into the intricacies of AI, starting with a brief history, followed by the various AI concepts such as machine learning and AI contracting applications, and concludes with a discussion on the acquisition of emerging technologies such as AI. The fourth section, "Methodologies," describes the three methodologies, EVM, KVA, and IRM in more detail, as well as provides examples of how they are applied. The final section, "Conclusion," wraps up the research with a series of conclusions and recommendations for future research.

This report seeks to introduce tried and true quantitative methods (EVM, KVA, IRM) for their application to the DoD Acquisition of AI systems. Therefore, this research recommends that future studies examine whether acquisition methods, strategies, and methodologies should change based on the category of AI being acquired.

THIS PAGE LEFT INTENTIONALLY BLANK

Table of Contents

THIS PAGE LEFT INTENTIONALLY BLANK

Acronyms

AC - Actual Cost

ACAT - Acquisition Category

ACWP - Actual Cost of Work Performed

AI - Artificial Intelligence

AoA - Analysis of Alternatives

APB - Acquisition Program Baseline

ARIMA - Autoregressive Integrated Moving Average

ARP - Acquisition Research Program

BCWS - Budget Cost for Work Scheduled

CAE - Component Acquisition Executive

C4ISR - Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance

CPI - Cost Performance Index

CV - Cost Variance

DAE - Defense Acquisition Executive

DAS - Defense Acquisition System

DoD - Department of Defense

EMD - Engineering and Manufacturing Development

EV - Earned Value

EVM - Earned Value Management

FAR - Federal Acquisition Regulation

GARCH - Generalized Autoregressive Conditional Heteroskedasticity

ICD - Initial Capabilities Document

IOC - Initial Operating Capability

IRM - Integrated Risk Management

IT - Information Technology

JCIDS – Joint Capabilities Integration and Development System

JTRS - Joint Tactical Radio System

KPP - Key Performance Parameter

KSA - Key System Attributes

KVA - Knowledge Value Added

LCSP - Life Cycle Sustainment Plan

LOC - Line of Communication

LRIP - Low Rate Initial Production

MAIS - Major Automated Information System

MBSE - Model-Based System Engineering

MSA - Materiel Solution Analysis

NDA - Non-Disclosure Agreement

NOAA - National Oceanic and Atmospheric Administration

NPS - Naval Postgraduate School

OS - Operations and Support

OT&E - Operational Testing and Evaluation

PD - Production and Deployment

PDR - Preliminary Design Review

PM – Program Manager

POA&M - Plan of Action & Milestones

PPBE - Planning, Programming, Budgeting, and Execution

PV – Planned Value

RDT&E - Research, Development, Test & Evaluation

RFP - Request for Proposals

RMF - Risk Management Framework

ROI - Return on Investment

ROK - Return on Knowledge

ROKI - Return on Knowledge Investment

SME - Subject Matter Expert

SOS – System of Systems

SV - Schedule Variance

TMRR - Technology Maturation and Risk Reduction

TRA - Technology Readiness Assessment

TRL - Technology Readiness Level

WBS - Work Breakout Structure

THIS PAGE LEFT INTENTIONALLY BLANK

Introduction

Acquisitions of artificial intelligence (AI) systems is a relatively new challenge for the Department of Defense (DoD). Given the high risk of failure for such system acquisitions, it is critical that the acquisition community examines potential new approaches to help manage the AI acquisition life cycle. The well-documented recurring problems in acquiring information technology within the DoD will likely be exacerbated in acquiring these leading edge, complex, and risky systems. The identification, review, and recommendation for the optimal use of new acquisition methodologies, to supplement or replace existing methodologies, should help avoid costly disasters in AI system acquisitions. In addition, the use of these methodologies should also create a more flexible acquisition scheme that allows for incorporation of unanticipated, value-added components of future AI systems.

AI has been in use in various commercial and governmental domains to address a variety of decision support problems. However, existing DoD acquisition frameworks may not be adequate to address the unique nature of AI systems life-cycle investments. AI systems are qualitatively different than standard automation systems that focus on routine, repeatable tasks. To develop acquisition frameworks for AI systems, it is first necessary to examine how AI systems will be used to support, or supplant, decision-makers. The purpose of this research project is to provide a set of quantitative and analytically robust decision-making methodologies for acquiring AI systems that address the inadequacies of the current standard investment acquisition life-cycle framework.

To better understand the potential contributions of this research, it is important to recognize the recent drive toward using innovation in improving defense acquisition outcomes. During the Cold War, the United States and particularly the Department of Defense enjoyed a position of prominence in the realm of military technological development. Therefore, the use of the common Defense Acquisition System allowed the DoD to develop, test, and field large-scale weapons systems through a slow, labor intensive development process. However, the rapid growth of technological developments has diminished the DoD’s technological advantage over many of the

United States' near peer global competitors. As a result of the current challenges, the DoD launched several key initiatives such as Better Buying Power 3.0, which is "aimed at accelerating acquisition reform and incentivizing innovation within government," as well as the Defense Innovation Initiative, and the "Third Offset Strategy" (Voelz, 2016, p. 180).

The underlying premise behind these recent efforts is that the traditional methods of acquisition are less than optimal in achieving the desired outcomes in DoD weapons and business systems. Therefore, the goal of implementing innovation in the DoD acquisition process has been designed to provide the means of achieving these better acquisition outcomes. Some of the key attributes in the recent initiatives include adopting best practices in DoD labs (Sullivan, 2018) and increasing the use of rapid development cycles through prototyping (DiNapoli, 2019). In addition, the DoD also seeks to take advantage of the commercial sector's rapid development of innovative technologies by partnering with the commercial development of dual use technology—technology that has application for both the private and defense sectors (Kendall, 2017; Voelz, 2016). It is this urgent context that provides a push for innovation in the DoD acquisition processes and provided the impetus for the current research study.

Research and Problem Statement

The current problem at the DoD is that the complexity and speed of decision-making is increasing exponentially with the advent of intelligent systems that support or that actually make decisions in time-critical, high-impact problem spaces. The current process management and control tools that a program manager (PM) might use to support acquisitions do not provide adequate warning of, or provide sufficient information about the root causes of, fiscal budgetary overruns and time schedule delays. This is a problem because PMs are, as a result, unable to respond to issues in a timely manner, delaying the delivery of promised capabilities to the services. Additionally, the money and resources spent in excess of the original budget could be used in other acquisition programs. To better understand the possible causes and solutions to the AI acquisition problem, this study examines the strengths and weaknesses of three performance and project management methodologies. These methodologies, Earned Value Management (EVM), Knowledge Value Added (KVA), and Integrated Risk Management (IRM), are used to strategically and tactically plan, monitor in real time, measure, and preemptively forecast the value and progress of AI acquisitions. A review of these recommended project analysis and control methodologies will offer insights into the strengths and weaknesses each approach could offer acquisition professionals within the general phases of the Defense Acquisition System. This research offers potential solutions to improve early warnings of cost and schedule overruns, and value opportunities foregone in the acquisition process. As such, this research focuses on the review of these methodologies and their applications to the acquisition process of AI systems.

Research Objective

The current research examines the challenges of acquiring an AI-based system within the common acquisitions framework in business and in the DoD. The primary objective of the current research is to evaluate three quantitative analysis tools for improving the acquisition of organically developed DoD AI systems. A comparison and contrast of the proposed methodologies will identify when and how each method can be applied to improve the acquisitions life cycle for AI systems.

Research Questions

The questions examined in this research are as follows:

1. How should each proposed methodology be used throughout the AI systems acquisition life cycle?
2. Will the combination of methodologies reduce the risks associated with acquiring an organically developed AI system?
3. When should the methodologies be used in the acquisition life cycle to ensure successful acquisition of AI systems?
4. What are the risks inherent in following a 5000-series acquisition framework when acquiring an organically developed AI system?

Technical Approaches and Outcomes of the Research

This research provides an in-depth review of each of the three methodologies (IRM, KVA, and EVM). While each acquisition project is unique, all must pass a series of common hurdles to succeed. A successful AI acquisition approach requires the support of methodologies that are designed to identify and value system options and forecast the future value of systems while assessing and mitigating investment risks. The dominant methodology for managing DoD acquisitions that exceed $20 million is EVM. The current structure of EVM may be enhanced with the addition of the IRM and KVA methodologies due to the unique needs of an AI system acquisition. This research examines how these three methodologies might be incorporated within an acquisition life-cycle framework assessing the benefits and risks of this potential extension of the standard framework. These methodologies have been used extensively in the past in acquisition research performed for the Acquisition Research Program (ARP) at the Naval Postgraduate School (NPS). This current study builds on the key learnings from this prior research to enhance the acquisition life-cycle framework with a focus on the unique characteristics of AI system acquisitions. The anticipated outcome of this research will be a set of guidelines for how and when to use the three methodologies to improve the potential success of acquisition of AI within the acquisition's life cycle.

The history of organic complex information technology (IT) has been characterized by cost and schedule overruns that have created havoc for acquisition professionals as well as systems designers and future users who expect to receive valuable new capabilities (Oakley, 2020; Housel et al., 2019a). The DoD's standard 5000-series acquisition life cycle provides the context for reviewing the ways the methodologies can be used to enhance the acquisition life-cycle approach in managing the acquisition of AI systems.

Risk-Tolerant and Risk-Averse Behavior in DoD Acquisitions

This following discussion of programmatic risk management requires a foundational understanding of the effect of risk on human behavior. Primarily this is the effect of risk on the PM and other program leaders. The goal is to gauge how these principal acquisition professionals respond to risk and their aversion to risk with regard to acquisition decision-making in terms of cost, schedule, and performance. Bhatt et al. (2005) note that a fundamental understanding of risk management addresses the question, "How much risk is acceptable?" (p. 64). As noted by Housel et al. (2019b), "A recurring issue at the U.S. Department of Defense (DoD) is that acquisitions of information technology (IT) have been fraught with schedule and cost overruns. The problem is the risk and project management tools the DoD currently uses inadequately address the fiscal and temporal overruns" (p. 3). This premise is supported by numerous studies on DoD acquisitions, particularly those that involve complex IT systems. A prime example that illustrates this issue is the multitude of reports on the development of space and satellite systems (Chaplain, 2017, 2019; Ludwigson, 2019). The issue of concern is that PMs and other managers of DoD acquisitions, particularly in the case of advancing cutting-edge technological systems, are increasingly becoming either overly risk tolerant or increasingly risk averse. As noted by Chaplain (2019) in the GAO report:

> Cost and schedule growth in DOD's space programs is sometimes driven by the inherent risks associated with developing complex space technology; however, over the past 10 years we have identified a number of other management and oversight problems that have worsened the situation. These include making overly optimistic cost and schedule estimates, pushing programs forward without sufficient knowledge about technology and design, and experiencing problems in overseeing and managing contractors, among others. (p. 6)

Thus, the motivation to be risk averse promotes a cautious culture among the DoD acquisition community. As the DoD continues to develop systems that are increasingly complex, the risk tolerance of PMs and other acquisition leaders is diminished in direct proportion to their inability to meet the program requirements for cost, schedule, and performance. It follows that the result of this is an organizational culture in the DoD acquisition community that is unwilling to assume risk. In addition, there is tremendous pressure to push unrealistic schedule and cost goals and results in unforced errors in the acquisition program's baseline. As a result of these unintended consequences, programs are most often more expensive, late, and/or not able to perform to the standards specified.

Integrated Risk Management

IRM is a comprehensive methodology that is a forward-looking risk-based decision support system incorporating various methods such as Monte Carlo Risk Simulation, Stochastic Forecasting, Portfolio Optimization, Strategic Flexibility Options, and Economic Business Case Modeling. Economic business cases using standard financial cash flows and cost estimates, as well as non-economic variables such as expected military value, strategic value, and other domain-specific subject matter expert (SME) metrics (e.g., Innovation Index, Conversion Capability, Ability to Meet Future Threats, Force Structure, Modernization and Technical Sophistication, Combat Readiness, Sustainability, Future Readiness to Meet Threats) can be incorporated (Mun, 2016a). These metrics can provide robust forecasts as well as mitigating risk via simulations that account for program uncertainties. The tools set also uses modeling to determine potential program benefits compared to program costs (e.g., return on investment for innovation or return on sustainability). Capital investment and acquisition decisions within AI program investment portfolios can then be made based on the resulting rigorous quantitative analysis (considering budgetary, manpower, and schedule constraints). Projects can be broken down into their detailed work breakout structure (WBS) and tasks, where these tasks can be combined in complex systems dynamic structures or implementation paths. The cost and schedule elements for each task can be modeled and risk-simulated within the system to estimate the resulting total cost and schedule risk of a given AI acquisition program. Portfolio management is often integrated with IRM methods to provide a more holistic view in terms of acquisitions of IT and AI acquisition programs.

Knowledge Value Added

KVA identifies the actual cost and value of an organization's assets (human and technological), standard functional areas, or core processes. KVA identifies every process required to produce an output, and the historical costs of those processes, the unit costs, and unit values of products, processes, functions, or services can be measured. By describing all process and subprocesses (down to the detailed level of WBS) outputs in common units, the methodology also permits market-comparable data to be generated; this ability is particularly important for nonprofits like the military and government organizations. Value is quantified in two key productivity metrics: Return on Knowledge (ROK) and Return on Investment (ROI) (when market comparables from industry are used). KVA includes the following seven-step method (Housel & Kanevsky, 2006):

- Identify functional areas and core processes along with their subprocesses. It is quite useful to have at least two process or functional area subject matter experts (SME) to ensure reliable estimates.

- Establish common units of output and level of aggregation of the process output to measure learning time required to produce the outputs. Other common-unit measures of output can also be used such as tasks (e.g., WBS), computer code, or process instructions that may be contained in existing documentation as long as they are calibrated to a common level of complexity using learning times.
- Calculate learning time (i.e., knowledge surrogate) required to produce the outputs of each process or functional area.
- Designate a sampling time period long enough to capture a representative sample of the core processes’ or functional area’s aggregated output.
- Multiply learning time for each process by the number of times the process executes during the sample period.
- Calculate the cost to execute the knowledge (e.g., learning time or process instructions) used by the resource (i.e., people, technology) used to produce the outputs to determine process costs.
- Calculate ROK and ROI.

Following these steps yields a defensible estimate of the productivity (i.e., ROK, ROI) of a given process or set of subprocesses. These estimates can then be used to track progress in an EVM framework in terms of cost, schedule, and importantly, the value produced. The KVA estimates can also be used to track the volatility of a set of processes, and this metric can be used in the IRM processes that forecast future value from, for example, an AI system.

Earned Value Management

EVM provides cost and schedule metrics to track performance in accordance with an acquisition project plan. EVM is required for large DoD acquisition programs that use incentive contracts valued at or greater than $20 million (Department of Defense [DoD], 2015a).

EVM methodology uses a WBS to try to ensure that an acquisition project is on schedule and within the estimated cost for each work package. It is used to measure work progress and any deficiencies using cost and schedule metrics that also can be used to measure program performance trend analysis with a focus on identifying any budget and schedule deviations from plan. However, the analysis is done after each process or stage in the WBS. In other words, the actual cost and time spent to execute a particular phase is compared against the initially projected budgeted plan.

Given the propensity of IT and AI acquisitions to be over budget and behind schedule, EVM metrics help PMs identify and attempt to avoid overruns and schedule deviations. Recognized plan deficiencies can help program managers identify waste and chokepoints that require immediate correction. When deficiencies in cost or schedule occur, EVM analysis can be used to reforecast the budget and schedule with the focus of providing PMs with up-to-date accurate performance information. EVM analysis uses schedule and cost estimates to find the Planned Value (PV) of a given acquisition project. Cumulative PV provides the total value that should be achieved by a specified date. The specific label for PV within the DoD acquisitions community is Budgeted Cost for Work Scheduled (BCWS). Actual Cost (AC) is the accumulated accrued costs of labor and materials at any point in time during a project. The label for AC within the DoD acquisitions community is Actual Cost of Work Performed (ACWP). Earned Value (EV) measures the progress for a given plan. The DoD acquisitions label for EV is Budgeted Cost of Work Performed (BCWP).

Research Report Layout

The next section provides a detailed list of the literature survey performed, with an emphasis on the Defense Acquisition System, followed by the various approaches: EVM, KVA, and IRM. The third section delves into the intricacies of AI, starting with a brief history, followed by the various AI concepts such as machine learning and AI contracting applications. The fourth section lists the methodologies in more detail, and provides examples of how they are applied. The final section wraps up the research with a series of conclusions and recommendations.

Literature Survey

This section starts with a discussion of the Defense Acquisition System (DAS) and the system acquisition life-cycle methodologies that might support the acquisition process. Each methodology is then described, providing a basic understanding of its purpose as well as how the method can be applied to the acquisition life cycle. This review provides a basic understanding of these methods whereas the next section covers each methodology in more detail.

Defense Acquisition System

The DoD oversees the acquisition of new systems through the Defense Acquisition System, which manages national investment in technologies, programs, and product support for the United States Armed Forces (Department of Defense [DoD], 2003). Its primary objective is “to acquire quality products that satisfy user needs with measurable improvements to mission capability and operational support, in a timely manner, and at a fair and reasonable price” (DoD, 2003). Within the DoD Decision Support System, there are three separate but interrelated processes: Joint Capabilities Integration and Development System (JCIDS), Planning, Programming, Budgeting, and Execution (PPBE) process, and the Defense Acquisition System (DoD, 2017). This research focuses on program management, as opposed to contract management, within the Defense Acquisition System.

Acquisition programs are divided into different ACATs based on the type of program and the dollar amount that is spent or is projected to be spent within the program (DoD, 2015a). Figure 1 shows the various cost-based designations and categories within the Defense Acquisition System. All dollar amounts for ACAT classifications are calculated in fiscal year 2014 dollars (DoD, 2015a). ACAT I is designated for major defense acquisition programs with an estimated Research, Development, Test & Evaluation (RDT&E) expenditure of more than $480 million or more than $2.79 billion for the total procurement (DoD, 2015a). An ACAT IA designation is for major automated information systems that will exceed $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).

ACAT II programs do not meet the criteria for ACAT I and will spend more than $835 million in the total procurement (DoD, 2015a) or more than $185 million in RDT&E. Finally, ACAT III programs are those not meeting the criteria for ACAT I or ACAT II designation (DoD, 2015a). The various designations allow for decentralized control of a program as each category has different reporting requirements and designated decision-makers (DoD, 2017).

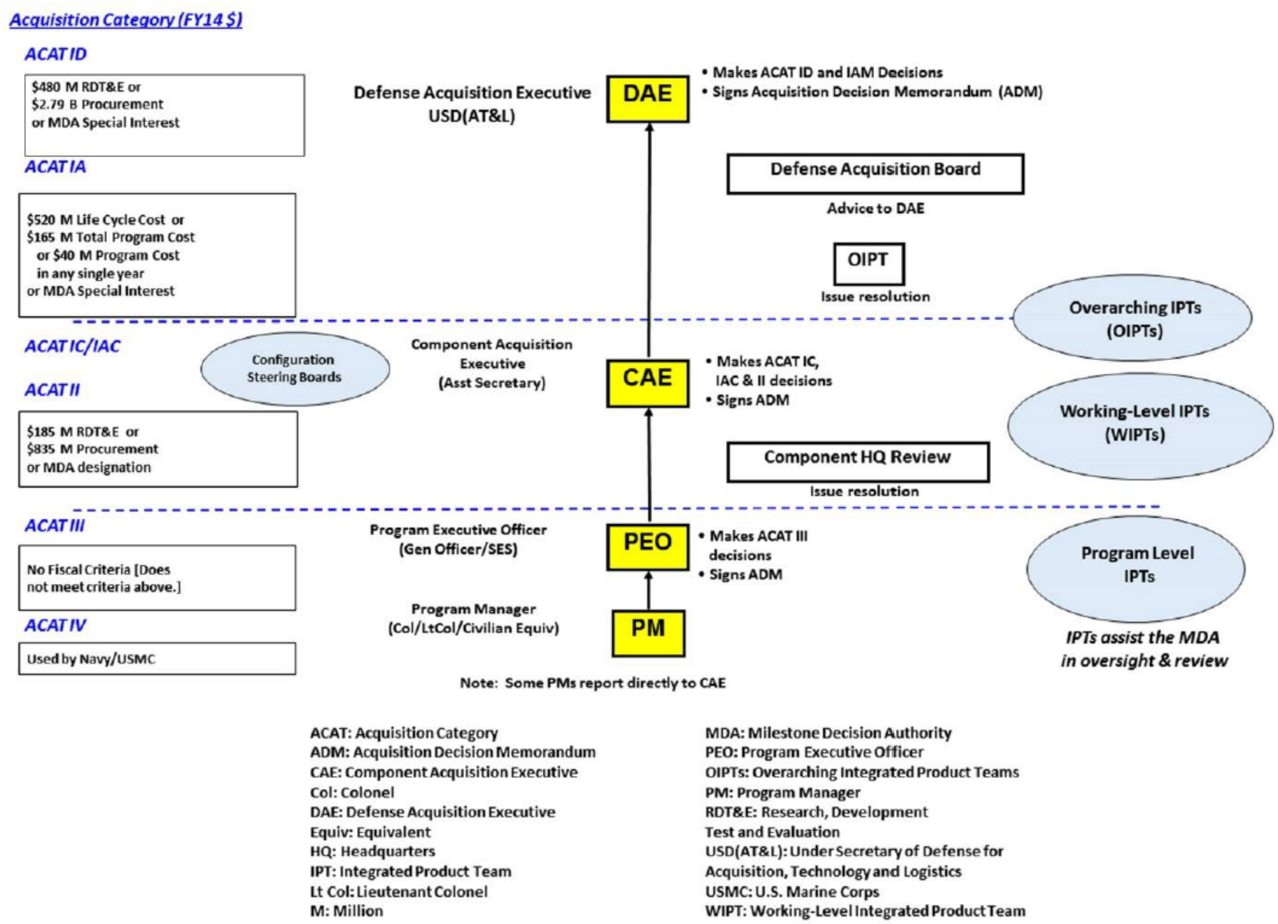


Figure 1. Acquisition categories

***Source:* DoD (2017)**

There are five phases within the Defense Acquisition System:

1. Materiel Solution Analysis (MSA)
2. Technology Maturation and Risk Reduction (TMRR)

3. Engineering and Manufacturing Development (EMD)
4. Production and Deployment (PD)
5. Operations and Support (OS)

Requirements for new or improved capabilities, delivered through the JCIDS process, drive the acquisition process (DoD, 2015a). Figure 2 illustrates the relationship between the acquisition and capabilities requirements processes and their interaction in the various acquisition phases. This study assumes the capabilities requested from the JCIDS process are accurate and necessary.

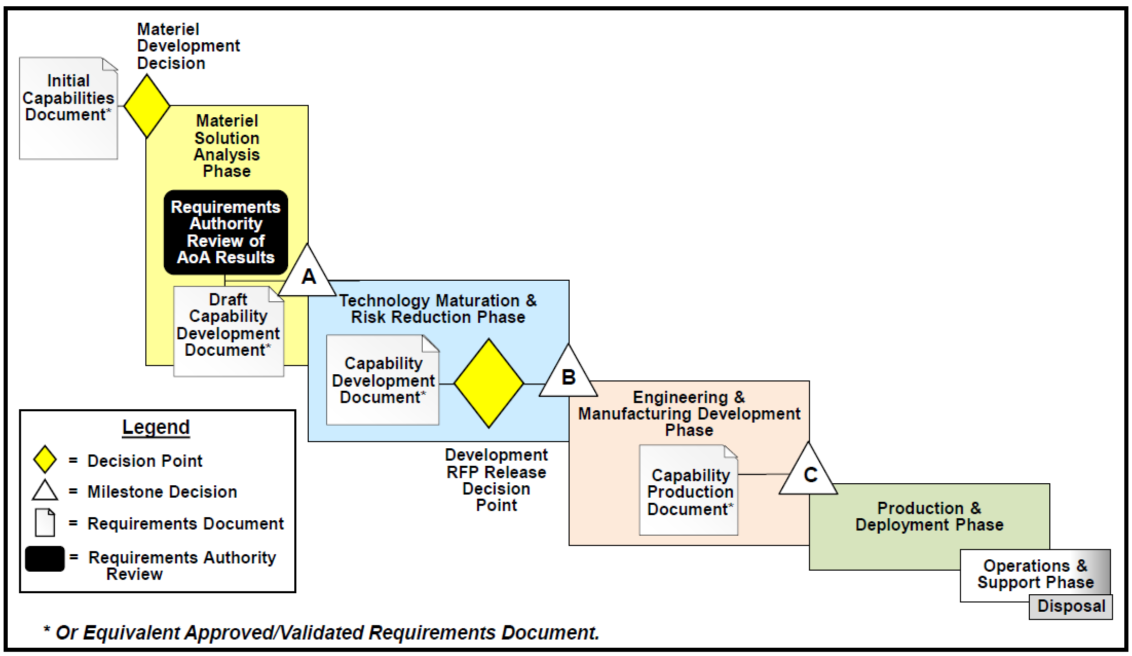


Figure 2. Interaction of capabilities requirements and acquisition process

Source: **DoD (2015a)**

Once an Initial Capabilities Document (ICD) has been validated, the Materiel Development Decision initiates the MSA phase (DoD, 2015a). This decision begins the acquisition process, although an acquisition program is not officially created until Milestone B at the completion of the phase (DoD, 2015a). The purpose of the MSA phase is to choose the most promising potential solution for the acquisition process that will fill the needs of the ICD and to establish Key Performance Parameters (KPPs) and Key System Attributes (KSAs) for the system (DoD, 2015a). To accomplish this, an Analysis of Alternatives (AoA) is conducted to determine the suitability of potential acquisitions based on “measures of effectiveness; key trades between cost and capability; total life-cycle cost, including sustainment; schedule; concept of operations; and overall risk” (DoD, 2015a, p. 17). The PM is selected, and the Program Office is established during this time (DoD, 2015a). Once the necessary analysis is concluded, 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 continue to the next phase based on the justification for the chosen solution, how affordable and feasible the solution is, how adequate the cost, schedule, and technical risk mitigation plan is, and how effective the acquisition strategy will be (DoD, 2015a). This decision is known as Milestone A (DoD, 2015a). The MSA phase takes a broad look at the potential solutions to a stated need and, as such, may be an appropriate place to consider strategic methodologies such as KVA or IRM.

After approval at Milestone A, the program enters the TMRR phase to reduce the risk associated with the technology, engineering, life-cycle cost, and integration of the program to begin the EMD phase (DoD, 2015a). Design and requirement trades occur at this point that are based on the budget, schedule, and likelihood of completion (DoD, 2015a). Guided by the acquisition strategy approved at Milestone A, contractors develop preliminary designs, including competitive prototypes if feasible within the program, to demonstrate the feasibility of their proposed solutions to the program office (DoD, 2015a).

Technology Readiness Levels (TRLs) serve as benchmarks that indicate the level of risk associated with a solution reaching maturation per the schedule (DoD, 2015a). Technology Readiness Assessments (TRAs) are a systemic, metric-based method to evaluate the maturity and risk associated with the critical technology in an acquisition program (DoD, 2011). A TRA will assign a TRL for each critical technology in a program, ranging from 1 to 9 from the lowest to highest readiness level (DoD, 2011). Additional methods to assess the likelihood a program will remain on schedule and on budget may be beneficial at this stage, such as IRM. The Development Request for Proposals (RFP) Release Decision Point authorizes the release of an RFP with firm and clearly stated program requirements for contractors to submit their bids (DoD, 2015a). The Preliminary Design Review (PDR) occurs prior to the completion of the TMRR phase unless waived by the milestone decision authority (DoD, 2015a). Milestone B approves a program to enter the EMD phase and awards a contract while establishing the Acquisition Program Baseline (APB; DoD, 2015a). The APB describes the approved program, specifically the cost and schedule for the life of the program and is a formal commitment to the milestone decision authority (DoD, 2015a).

EMD begins once Milestone B is approved. During EMD, the materiel solution is developed, built, and tested to verify all requirements have been met prior to production (DoD, 2015a). Hardware and software designs are completed, and prototypes are built to identify any deficiencies in the design, which will be discovered during developmental and operational testing (DoD, 2015a). DoD acquisitions programs with a contract value greater than $20 million are required by federal regulation to use EVM to track and report the progress of the program, which begins during this phase (DoD, 2019a). Once a stable design that meets the specified requirements has been verified, the manufacturing or software sustainment processes and production capability must be properly demonstrated (DoD, 2015a). Milestone C confirms that these requirements are satisfied and approves entry into the PD phase (DoD, 2015a).

The objective of the PD phase is to deliver a product that fulfills the requirements specified in the earlier stages (DoD, 2015a). Initial operational deployment and testing occurs with Low Rate Initial Production (LRIP) for manufactured systems or limited deployment for more software-intensive programs where the system undergoes Operational Test[ing] and Evaluation (OT&E) to verify that stated requirements were met (DoD, 2015a). Once satisfied with the fielded systems, full-rate production begins, and the product is deployed to operational units (DoD, 2015a). Design changes are limited at this point, although some changes may still occur based on noted deficiencies (Housel et al., 2019a). Contracts typically revert to a fixed price strategy during this phase, reducing the PM's focus on cost and schedule variance (Housel et al., 2019b).

OS is designed to maintain support for the product and sustain its performance throughout its life cycle, ending with the disposal of the system (DoD, 2015a). OS overlaps with the PD phase since operational units are using the product while production continues, beginning after the production or deployment decision (DoD, 2015a). PMs will sustain the system using the Life Cycle Sustainment Plan (LCSP) developed during the acquisition process, providing the necessary resources and support to keep the system operational (DoD, 2015a). Sustainment and support may include technological upgrades, changes due to operational needs, process improvements, and other activities that may require updates to the LCSP (DoD, 2015a).

There are six different models—four are standard and two are hybrid—that PMs use to create their program structure, depending on the type of system being acquired (DoD, 2015a). These standard models are templates for hardware-intensive programs, software-intensive programs that are defense unique, incrementally deployed software-intensive programs, and accelerated acquisition programs (DoD, 2015a). As shown in Figure 3, the hybrid models mix the incremental nature of software development within a hardware-centric program. In this model, software development is organized via a series of testable software builds that will culminate with the fully required capability before reaching the Initial Operating Capability (IOC; DoD, 2015a). The incremental builds are synchronized with hardware testing requirements for prototypes and other developmental requirements (DoD, 2015a). Other models use the same basic framework within the five phases, with the exception of the accelerated program.

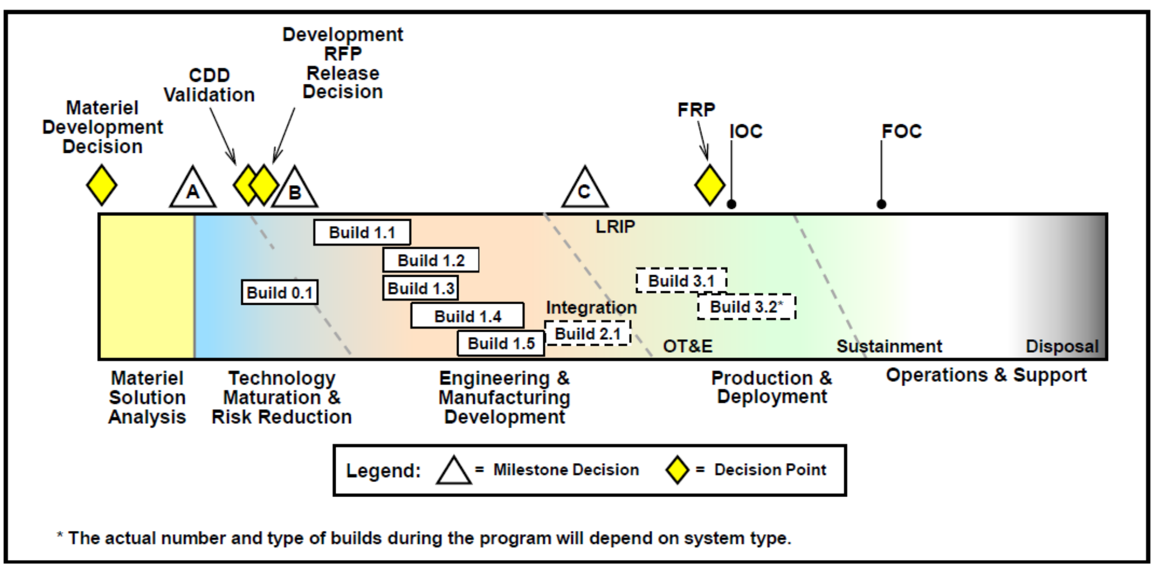


Figure 3. Hardware-dominant hybrid program

***Source:* DoD (2015a)**

AI and IT systems are increasingly prevalent throughout the DoD along with their connection to weapon systems, facilities, and Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR; DoD, 2015b). With the integration comes an increased security risk from adversaries, elevating the importance of effective cybersecurity capabilities and practices (DoD, 2015b). The DoD manages cybersecurity policy through the Risk Management Framework (RMF) by applying security controls founded on risk assessments throughout the life cycle of a system (DoD, 2015b). RMF applies to "all DoD IT that receive, process, store, display, or transmit DoD information" (DoD, 2014, p. 2). Cybersecurity within RMF is more than simply information security, including items such as stable and secure engineering designs, training and awareness for all users, maintainers, and operators of a program, and the response, recovery, and restoration of a system following an internal or external failure or attack (DoD, 2015b). The RMF occurs throughout the acquisition process, and Figure 4 illustrates the six steps of its process.

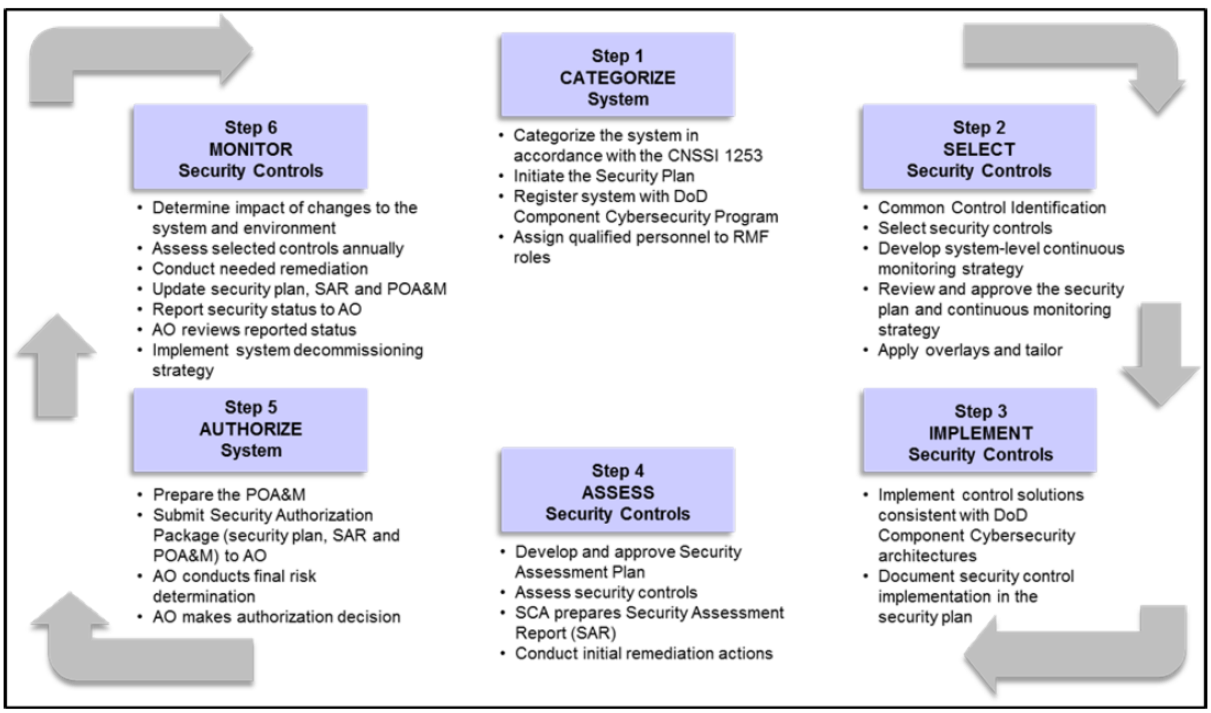


Figure 4. Risk management framework process

Source: **DoD (2014)**

The first step is categorizing the system, during which the potential impact of a breach is analyzed, and the system and its boundaries are described (DoD, 2014). The security plan is initiated, the system is registered with the DoD Component Cybersecurity Program, and the RMF team is formed (DoD, 2014). Cybersecurity requirements are included in the ICD, driving considerations in the MSA phase during the AoA (DoD, 2015b). The risk assessment considers the potential impacts on missions resulting from a cybersecurity breach (DoD, 2015b). The RMF provides a relatively objective method to determine the cybersecurity risk level that establishes the initial baseline security controls necessary to ensure they are included in the acquisition plan for the system (DoD, 2015b).

In step two, the RMF team selects security controls, including those controls common to other DoD programs (DoD, 2014). A plan to continuously monitor the effectiveness of the controls is developed and documented (DoD, 2014). This plan is then submitted to the DoD Components that review and approve the security plan (DoD, 2014). As the cybersecurity strategy is developed in the MSA phase, the acquisition and cybersecurity teams coordinate to ensure the appropriate level of security is implemented in the program throughout its life cycle and in the system architecture and design (DoD, 2015b). The continuous monitoring strategy and security plan are also developed during MSA (DoD, 2015b).

Next, the approved security controls are implemented per DoD guidelines (DoD, 2014). The implementation must be appropriately documented in the system's security plan (DoD, 2014). Cybersecurity requirements are part of the system performance requirements in the TMRR phase (DoD, 2015b).

Then, the RMF team must develop, review, and approve a Security Assessment Plan that will allow proper assessment of the security controls (DoD, 2014). Once approved, the system security is assessed in accordance with DoD assessment procedures and the Security Assessment Plan, during which vulnerabilities are assigned severity values and the security risk for both the controls and the aggregate system is determined (DoD, 2014). This is documented in the Security Assessment Report, which is required before authorization of any system, and remediation actions on the security controls are conducted (DoD, 2014). The cybersecurity requirements stated in the

Capability Development Document are validated during the TMRR phase prior to a RFP (DoD, 2015b). The Preliminary Design Review, which is also conducted during the TMRR phase will include cybersecurity aspects, ensuring the approved plan is implemented in the chosen design and risks are mitigated to an appropriate level (DoD, 2015b). As the system develops in the EMD phase, all computer code follows applicable standards and secure coding practices with assessments conducted and documented in the Security Plan (DoD, 2015b).

Based on the recognized vulnerabilities, a Plan of Action and Milestones (POA&M) is created that identifies tasks needed to mitigate the vulnerabilities, resources necessary to complete the plan, and milestones toward completing tasks (DoD, 2014). The Authorizing Official who will determine if the risk level is appropriate prior to authorizing the system receives the Security Authorization Package (DoD, 2014). Creation of the POA&M begins in the MSA phase and continues throughout system development (DoD, 2015b).

Finally, the security controls must be monitored throughout the life of the system to ensure any changes to the system or the environment do not negatively affect cybersecurity measures (DoD, 2014). Should someone detect vulnerabilities, the necessary remediation will be conducted, and the security plan updated (DoD, 2014). Once a system is approved and operationally deployed, the cybersecurity is monitored in accordance with the continuous monitoring strategy and Security Plan (DoD, 2015b). New risk assessments are conducted when changes to the system, its environment, or the planned use of the system occur (DoD, 2015b). Should vulnerabilities occur, the PM updates the Security Plan and POA&M to indicate how the vulnerability will be addressed (DoD, 2015b).

Earned Value Management

Currently the DoD uses EVM within the Defense Acquisition System (DAS) to evaluate the progress of acquisitions. EVM is a system PMs use to integrate the work scope with the cost and schedule of that program to improve the control and planning of the acquisition. It establishes a baseline for the objectives of the program to measure cost and schedule performance while the project is being executed. EVM is used to identify problems, create corrective actions to fix those problems, and allow management to re-plan the program as required (Electronic Industries Alliance, 1998). The Federal Acquisition Regulation (FAR) specifies that DoD acquisition programs whose contract value exceeds $20 million are required to use EVM in the program office (DoD, 2019). Mandates within the federal government require reports on the progress and execution of acquisition projects, leading to an emphasis on performance measures.

In sum, EVM exists to provide an assessment of the actual physical work a project has completed compared to a baseline plan (Fleming & Koppelman, 2010). EVM integrates the actual cost spent on the project to date with the work that has been performed on the project, allowing managers to compare the progress of the project with their planned budget and schedule (Fleming & Koppelman, 2010). It provides managers the ability to compare cost performance with work completion rather than simply cost performance and planned cost, as is done in traditional cost management (Fleming & Koppelman, 2010). When properly employed, EVM provides a reliable prediction of the total cost and schedule requirements for a project through three distinct dimensions: the PV, EV, and actual cost (Fleming & Koppelman, 2010).

PV, referred to within the DoD as Budgeted Cost of Work Scheduled (BCWS), is the amount of work, either physical or intellectual, scheduled to be completed by a certain point (Fleming & Koppelman, 2010). It is a time-phased budget reference and is used throughout the project as a baseline for the amount of work completed by the scheduled date (Vanhoucke, 2014). When depicted graphically (as in Figure 6) it is an upward-sloping function and shows the cumulative increase in all scheduled and budgeted activities from the beginning of the project until completion (Vanhoucke, 2014). Simply stated, BCWS is the authorized budget for authorized work (Fleming & Koppelman, 2010). This baseline should be established prior to a program’s initiation and should remain constant throughout the program to maintain a fixed reference, although the baseline can be re-established if performance is drastically different than originally planned to improve future project control (Vanhoucke, 2014).

To establish a baseline, the scope of a project must be fully defined, the resources necessary to complete the project must be understood, and the compulsory tasks must be placed into the timeline required to complete each task (Fleming & Koppelman, 2010). "If you do not know what constitutes 100% of a project, how will you ever know if you are 10, 20, or 35 percent done?" (Fleming & Koppelman, 2010, p. 48). Project managers create a work breakdown structure (WBS) to produce an accurate baseline. A WBS is a division of tasks arranged in a hierarchical, tiered fashion portraying the breakdown of activities used to authorize, track, and report a program's progress. It relates the individual elements necessary to complete work to each other and the system as a whole (DoD, 2005). A WBS can be expressed in any level of detail, from high-level systems view, such as Figure 5, down to the distinct pieces of material needed to construct a component, depending on the level of detail needed (DoD, 2005). Within the 5000 series, the BCWS baseline is usually established during the TMRR phase.

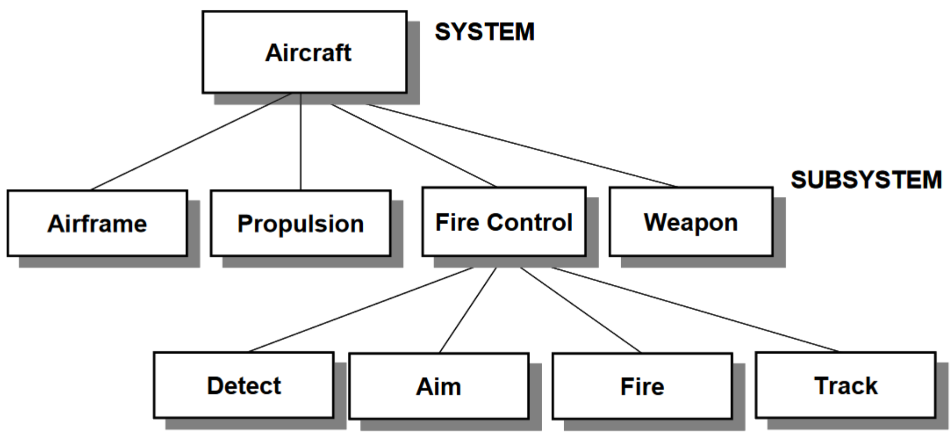


Figure 5. Sample WBS

EV, the second dimension within EVM, represents the amount of money from a project's total budget spent on the work accomplished at a certain point in time (Vanhoucke, 2014). Also referred to as the Budgeted Cost of Work Performed (BCWP), it shows the total budget of the completed work packages and finished sections of open work packages (DoD, 2019). BWCP comprises the amount of authorized work that was actually completed with the amount of the original budget for accomplishing the given work (Fleming & Koppelman, 2010).

The third dimension of EVM is actual cost (AC), or the Actual Cost of Work Performed (ACWP). ACWP is the cumulative total cost a program has spent to accomplish work at a given point in time (Vanhoucke, 2014). It measures the amount of money used to convert the PV into EV within the measured time frame (Fleming & Koppelman, 2010). ACWP depicts the amount of money spent on a project regardless of the output of the work. It is purely a financial metric illustrated over the elapsed time of a project and does not actually account for the work that is actually accomplished.

Figure 6 gives a graphical depiction of PV (BCWS), EV (BCWP), and AC (ACWP) for a fictitious project. In blue is the PV, showing the amount of money budgeted to complete specific work packages based on the WBS. Green displays the budgeted cost of the work packages that have been completed at a specific time, or EV. At the project's completion, EV and PV are equal since EV is calculated as a percentage of the planned budget. AC, shown in red, portrays the money spent to complete the EV at the same point in time. Ideally, all three lines will overlap, indicating the project is exactly on schedule and budget. However, this is rarely the case, and the differences indicate the need for additional information to determine what corrections are necessary, leading to the performance metrics.

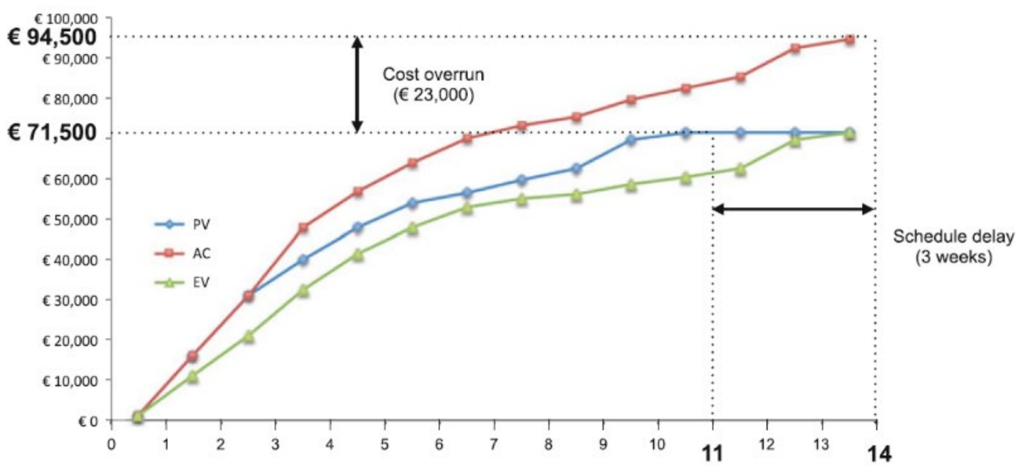
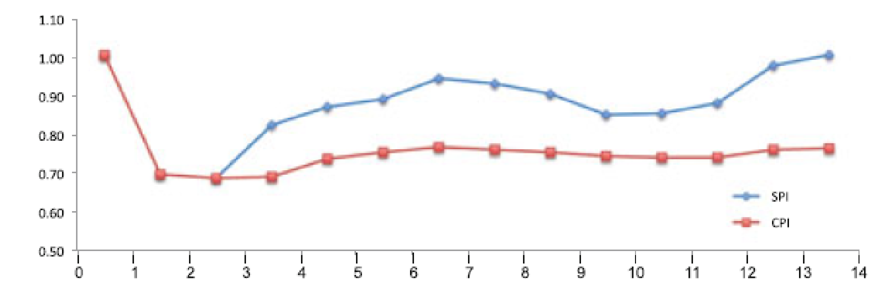


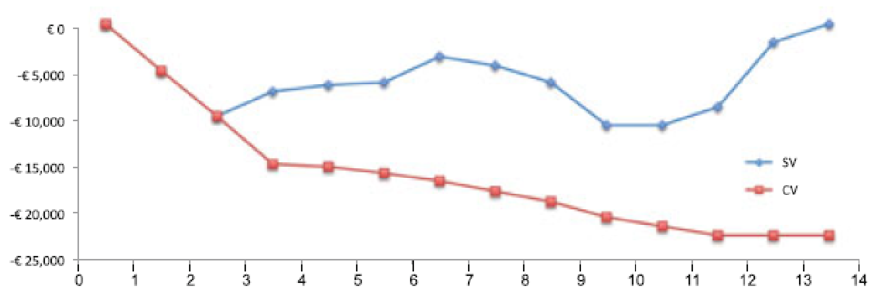
Figure 6. S-curve with the three EVM dimensions
Source: **Vanhoucke (2014)**

Four performance metrics within EVM provide indications of a program's current performance compared to the baseline cost variance (CV), cost performance index (CPI), schedule variance (SV), and schedule performance index (SPI; DoD, 2019). CV determines the difference between the EV work completed and the AC: CV = EV – AC

(Fleming & Koppelman, 2010). If the difference is less than zero, the project is over budget, greater than zero is under budget, and if equal to zero, the project is on budget (Vanhoucke, 2014). The CPI is the ratio of completed work to the budget, calculated by dividing EV by AC: CPI = EV/AC (Fleming & Koppelman, 2010). CPI can be used to forecast a range of total costs to finish a project based on the performance of the project to date (Fleming & Koppelman, 2010). If the CPI is greater than 1, the project is under budget, less than 1 is over budget, and if equal to 1, the project is on budget (Vanhoucke, 2014). Both CV and CPI measure the deviation in the value of the completed work (EV) and the cost of the work (AC; Vanhoucke, 2014). Figure 7 shows the performance metrics from the example project in Figure 6 with CV and CPI in red. The CPI drops to roughly 0.7 in just over a week before maintaining a relatively constant level, indicating the project is over budget, while the CV continues to become increasingly negative, showing the increasing amount of money spent above what was budgeted (Vanhoucke, 2014). Although the magnitude of the CV continued to increase, the CPI remained constant, denoting the project continued to earn value at 70% of the planned rate.



SPI	1.00	0.69	0.68	0.82	0.86	0.88	0.94	0.93	0.90	0.84	0.85	0.87	0.97	1.00
CPI	1.00	0.69	0.68	0.68	0.73	0.75	0.76	0.75	0.74	0.74	0.73	0.73	0.75	0.76



SV	0	-5,000	-10,000	-7,325	-6,570	-6,315	-3,560	-4,447	-6,333	-11,000	-11,000	-9,000	-2,000	0
CV	0	-5,000	-10,000	-15,225	-15,470	-16,215	-16,960	-18,147	-19,333	-21,000	-22,000	-23,000	-23,000	-23,000

Figure 7. Example performance metric curves

Source: **Vanhoucke (2014)**

Similarly, SV and SPI compare the performance of a project with respect to its planned schedule. In the same manner that CV and CPI examine cost, these metrics quantify the divergence in the value of the completed work (EV), and the amount of value expected at a given point in time (PV; Vanhoucke, 2014). SV is the difference between the EV work completed and the PV: SV = EV – PV (Fleming & Koppelman, 2010). If the difference is less than zero, the project is behind of schedule, greater than zero is ahead of schedule, and if equal to zero, the project is on schedule (Vanhoucke, 2014). SPI is the ratio of completed work to the scheduled time that work was completed, calculated by dividing EV by PV: SPI = EV/PV (Fleming & Koppelman, 2010). This ratio can be used to estimate the project completion date (Fleming & Koppelman, 2010). If the SPI is greater than 1, the project is behind schedule, less than 1 is ahead of schedule, and if equal to 1, the project is on schedule (Vanhoucke, 2014). Referring again to Figure 7, the SPI and SV for the previous project are shown in blue. The SPI initially dips to roughly 0.7 before climbing back to 1 at the end of the timeline, while SV varies in a correlated curve until increasing back to 0 at the completion of the project (Vanhoucke, 2014). This indicates a slower start to the project and a recovery toward the schedule as work proceeds, even though SV never equals 0 and SPI never equals 1—the corresponding values for on-schedule performance—until the conclusion. While it may not be initially evident, this tells PMs the program did not finish within the planned timeline.

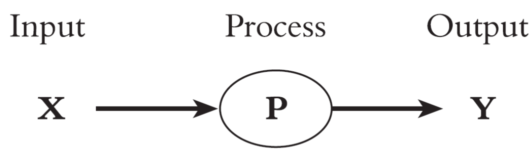
It is important to note the term *value* in EVM does not have the same meaning as in other methodologies, such as KVA. Within the context of EVM, *value* is defined as the work accomplished toward completion of the project. There is no reference to the quality of the completed work or additional (or missing) benefits the work might provide to a system. The value is assumed because the specifications were defined in the project requirements.

EVM has proven to be a reliable system to manage cost and schedule performance for manufacturing in both defense and commercial industries. However, as systems become more complicated and information technology (IT) and AI gains a more prominent place within even traditional manufacturing projects, EVM may need additional information from additional methodologies to improve its capabilities. Better incorporating the strategic guidance associated with a program, the value gained from subcomponents and subprocesses, the risk associated with developing subcomponents of a system, and incrementally improving a process may help improve the Defense Acquisition System as a whole.

Knowledge Value Added

KVA is an empirical model that focuses on the practical application and implementation of knowledge management (Tsai, 2014). Originally developed to assist in business process reengineering, KVA creates an objective, quantifiable method to measure the value of a process or service (Housel & Kanevsky, 1995). Typical financial approaches to business process reengineering use the dollar amount of a final product to determine the value of an object, failing to account for the knowledge required in the various subprocesses involved in making the product (Housel & Kanevsky, 1995). In its essence, KVA performs a single function: describing all process outputs in common units. KVA accounts for the value of all components, processes, and support systems necessary to complete a task or create a product or service by describing all outputs in common units. It allows managers to compare the efficiency of the various steps across all processes within a common value reference point.

Value has a different meaning in KVA than it does in other methodologies, such as EVM or IRM. KVA bases its definition of *value* on complexity theory and views organizational processes by their ability to change their input (raw material, information, energy, etc.) into common units of output, as shown in Figure 8 (Housel & Kanevsky, 1995). Per Figure 8, process P changes the input in some manner, creating a different product or service at the output, adding value to the system based on the number of common unit changes from input to output (Housel & Kanevsky, 1995). If process P did not change input X, then output Y is the same as input X, indicating no value was added by the system (Housel & Kanevsky, 1995). While the change from X to Y may be minute or large depending on the process, KVA converts all changes into common units, and these changes indicate the amount of value added by process P to produce the final product. The value generated through the process is proportional to the change in the state from X to Y, denoting the amount of knowledge required to make the changes (Yu et al., 2009). Thus, the contribution to a process is equivalent to the sum of all knowledge necessary to produce a product and/or interpret meaning from an input (Housel & Kanevsky, 2006). This is true for all processes within a system, from production to service to management.



P(X) = Y

Fundamental assumptions:

1. If X = Y no value has been added.

2. "value" ∝ "change"

3. "change" can be measured by the amount of knowledge required to make the change.

So "value"∝ "change" ∝"amount of knowledge required to make the change"

Figure 8. Value added process

***Source:* Housel & Bell (2001, p. 94)**

The KVA methodology is best completed by following the seven-step process shown in Figure 9. Practitioners can use several methods to describe the units of change, such as tasks, Haye knowledge points, Shannon bits, units of knowledge, and so on (Housel & Bell, 2001). For ease of measurement, three measures are typically used within KVA to estimate the embedded knowledge within a process (Housel & Bell, 2001). Learning time, column two in Figure 9, measures the length of time it takes an average user to learn a process and correctly complete it (Housel & Bell, 2001). Process description, column three, is the number of process instructions used to transform the given input into the desired output (Housel & Bell, 2001). Each instruction must require an approximately equal amount of knowledge to complete a task (Housel & Bell, 2001). The binary query method uses the number of binary questions (i.e., bits) necessary to accomplish the process, roughly equivalent to the lines of code within a computer program (Housel & Bell, 2001). However, any measure that satisfies the basic concepts of KVA can be used to create a common-units measure (Housel & Bell, 2001).

Steps	Learning time	Process description	Binary query method
1.	Identify core process and its subprocesses.		
2.	Establish common units to measure learning time.	Describe the products in terms of the instructions required to reproduce them and select unit of process description.	Create a set of binary yes/no questions such that all possible outputs are represented as a sequence of yes/no answers.
3.	Calculate learning time to execute each subprocess.	Calculate number of process instructions pertaining to each subprocess.	Calculate length of sequence of yes/no answers for each subprocess.
4.	Designate sampling time period long enough to capture a representative sample of the core process's final product/service output.		
5.	Multiply the learning time for each subprocess by the number of times the subprocess executes during sample period.	Multiply the number of process instructions used to describe each subprocess by the number of times the subprocess executes during sample period.	Multiply the length of the yes/no string for each subprocess by the number of times this subprocess executes during sample period.
6.	Allocate revenue to subprocesses in proportion to the quantities generated by step 5 and calculate costs for each subprocess.		
7.	Calculate ROK, and interpret the results.		

Figure 9. The KVA approach

***Source:* Housel & Bell (2001)**

The first step, regardless of which metric an analyst employs, is identifying the core process and its subprocesses (Housel & Bell, 2001). To fully understand and accurately measure the knowledge inherent in a process, the entirety of the process must be mapped (via a mapping of process inputs, process operation, and process outputs). Next, analysts determine the measure that will be used in the analysis to describe the subprocess outputs in common units (Housel & Bell, 2001). Learning time or lines of code (WBS can be used to represent the presumed output of a given step in the acquisition process) are the commonly used units as they can be used to convert outputs into common units relatively quickly depending on the degree of accuracy required. Analysts must then calculate the number of units (i.e., learning time, tasks, or lines of code) within each subprocess (Housel & Bell, 2001). Then, the actual measurement of output occurs over a specified period of time (Housel & Bell, 2001). The sample period will vary from system to system depending on the complexity and the length of each process. After determining the output for a standard execution of the process and the corresponding unit of measure (i.e., learning time, tasks, or lines of code) is established, the output is multiplied by the number of times each subprocess is used during the sample period (Housel & Bell, 2001). Next, a proportion of revenue (i.e., in for-profit organizations) is allocated to each of the subprocesses, by the number of common units of output produced by each subprocess, and costs are calculated for each subprocess (Housel & Bell, 2001). In the case of not-for-profits (e.g., DoD organizations), a market comparable aggregate revenue estimate can be calculated that provides a means to establish a price or revenue per common unit of output which is a constant as all common units of outputs are the same. Finally, analysts should determine the Return on Knowledge (ROK: monetized or non-monetized output divided by cost) and interpret the results (Housel & Bell, 2001). Analysts should use two or more estimates of description of the common unit of output based on the method selected for describing the unit of output (e.g., learning time, lines of code). These estimates can then be correlated to determine the reliability of the estimates. Estimates with a resulting high correlation ensure the reliability of the value calculations (Housel & Bell, 2001).

ROK is a ratio used to determine the value added from knowledge assets used within the project to produce outputs (Housel & Bell, 2001). It is calculated by dividing the knowledge embedded within a process and its frequency of use by the cost associated with operating that process (Housel & Bell, 2001). ROK can be calculated for any manual or automated activity; AI system and even management activities can be observed and measured via the KVA approach, due to the knowledge embedded in all of these resources. A higher ROK indicates more value returned for each dollar spent on the process (Housel & Bell, 2001). ROK gives acquisition managers a common reference point, an objective way to examine the benefit and value of a process compared to other processes or functions, allowing leadership to manage their processes within a portfolio framework to determine which, if any, process might benefit from moving knowledge from employees to automation (i.e., artificial intelligence, robotics, online applications) to make substantial improvements in productivity.

Figure 10 shows a rudimentary analysis of maintenance actions within a Marine Corps Motor Transport platoon via the learning time method previously discussed. The exercise was conducted to ascertain if an AI system would improve the timeliness of maintenance procedures while maintaining the same value in the process (Carlton et al., 2019). Total knowledge was calculated by establishing the total amount of formal training required to complete a given task, including initial and recurring training (Carlton et al., 2019). Expenses were estimated by multiplying the average salary of the Marine performing the work and the average time to complete the task (Carlton et al., 2019). The ROK was computed by dividing the total knowledge (i.e., translated into amount of output in common units) by the expenses (the cost component of the productivity ratio) (Carlton et al., 2019). After establishing the baseline as-is measurements, the team estimated the time necessary to complete each task (Carlton et al., 2019). Some steps within the process were automated using the proposed IT system, reducing the cycle time. However, total knowledge remained constant since the output from both the as-is and to-be processes were equal. ROK from the to-be process increased compared to the as-is process, suggesting the proposed AI solution will improve performance by reducing cost (Carlton et al., 2019). Since the total amount of output remained the same, this change suggests that cost will be reduced without destroying process output value. A more detailed example is presented in the next section.

KVA is potentially useful tool for inclusion in the Defense Acquisition System. Since the DoD is not a for-profit company, it does not have revenue to judge the effectiveness of its programs in a monetized form. Instead, it relies on various metrics and evaluations that are not comparable from system to system. If the DoD implements the KVA methodology more widely, PMs may have a more objective measure to compare various technological solutions to fulfill evolving requirements. Understanding the value that a system or process provides in direct comparison with the value of other systems, whether they are similar or unrelated processes, could provide beneficial information in the decision-making, budgeting, and planning processes.

As-Is Process

Process Description	Total Knowledge	Expenses ($)	ROK
Open SR	4320.00	$ 1,663.70	260%
Induction	87360.00	$ 3,327.39	2625%
Order Parts	10080.00	$ 332.74	3029%
Supply Opens Parts Requisition	20160.00	$ 1,663.70	1212%
Receive Part	41040.00	$ 1,996.43	2056%
Perform Maintenance	109920.00	$ 13,309.57	826%
Final Inspection	87360.00	$ 3,327.39	2625%
Totals	**360240.00**	**$ 25,620.91**	**1406%**

To-Be Process

Process Description	Total LT (Hrs)	Expenses ($)	ROK	ROK w/ Acq Costs
Open SR/Induction/OrderParts	101760.00	$ 3,992.87	2549%	2398%
Supply Opens Parts Requisition	20160.00	$ 1,663.70	1212%	1053%
Receive Part	41040.00	$ 1,996.43	2056%	1827%
Perform Maintenance	109920.00	$ 13,309.57	826%	811%
Final Inspection	87360.00	$ 3,327.39	2625%	2442%
Totals	**360240.00**	**$ 24,289.96**	**1483%**	**1468%**

Figure 10. Sample KVA tables

***Source:* Carlton et al. (2019)**

Integrated Risk Management

Risk-Tolerant and Risk-Averse Behavior in DoD Acquisitions

As noted in the introduction of this report, programmatic risk management requires that the PM understands the effect of one's own risk tolerance and risk aversion in the execution of an acquisition program. Conversely, often the PM and other Acquisition personnel impact their program adversely through changes in cost and schedule, and/or lack of performance through their own inherent aggressive risk tolerance or excessive risk aversion. As stated by Bhatt et al. (2005), a fundamental understanding of risk management is addressing the question, "How much risk is acceptable?" (p. 64). IT Acquisition Programs tend to be "fraught with schedule and cost overruns. The problem is the risk and project management tools the DoD currently used inadequately address the fiscal and temporal overruns" (Housel et al., 2019a). Along with IT acquisition systems, other DoD complex acquisitions, such as space and satellite systems also suffer the same challenge (Chaplain, 2017, 2019; Ludwigson, 2019). The issue of concern is that PMs and other managers of DoD acquisitions, particularly in the case of advancing cutting-edge technological systems, are increasingly becoming either overly risk tolerant or increasingly risk averse. However, there are quantitative tools that can reduce the impact of human factor of risk aversion and risk tolerance. One such method is for the PM to utilize modeling tools through processes such as IRM.

Model-Based Systems Engineering

Prior to reviewing the IRM methodology, it is important to understand the role that Model-Based Systems Engineering (MBSE) methodologies play in Defense Acquisition. Due to the nature of the DoD system development process including the procurements of complex systems, MBSE has been adopted as the common engineering approach when integrating work between diverse teams working on these complicated systems (Ramos et al., 2012). The increasing intricate complexity of these systems led to their characterization as the "System of Systems" (SoS) approach. SoS was developed to "describe systems that are composed of independent constituent systems, which act jointly toward a common goal through the synergism between them" (Nielsen et al., 2015, p. 18:1). The traditional means of systems engineers coordinating their multitude of efforts was through the back and forth exchange of numerous documentation artifacts related to the system or systems. Thus, Model-Based System Engineering (MBSE) is an approach that was established to reduce the complexity of the information exchange to "reduce development complexity, enhanced productivity, efficient change … for the development of embedded systems" (Rashid et al., 2015, p. 150).

The value of MBSE in defense acquisition is that it provides the DoD acquisition personnel in the fields of both systems engineering and program management the means to break down complex data sets and to share with other members on a team—regardless of location or role (i.e., government, contractor). Those involved in complex systems or SoS have made extensive use of MBSE as it motivates the numerous members of a diverse workforce to "speak" the same language, i.e., the model (Ramos et al., 2012). In the realm of PM modeling, the DoD has found MBSE useful for breaking down the components of system architectures such as the Department of Defense

Architectural Framework (DoDAF) from the traditional written architecture documentation into models that simplify terminology through basic syntax and symbology (Giachetti, 2015). Furthermore, the DoD has found exquisite simplicity to be beneficial in simplifying the components comprising a weapon system into MBSE symbols (Shin et al., 2017).

Cost Estimation

The acquisition decision process can be augmented with advanced data analytics. This is particularly true in the case of cost estimation. Acquisition cost estimation is comprised of two major approaches, quantitative and qualitative. In the case of quantitative there are four approaches to cost estimation, Analogy, Parametric, Engineering Build-up, and Extrapolation from Actuals (Lee, 2014). Analogy is used early in the process when limited data is available and the estimate must be made with similar items—as in the case of real estate estimates with comparable market values of homes (e.g., based on price per square foot). As more data becomes available, parametric modeling tools provide a stronger analytical model. And as more details emerge and more information is available for the cost estimator, a more accurate, Build-up cost estimation is used (Mun, 2019). Finally, when actual costs are identified, costs can be extrapolated, and thus, the cost estimation becomes the actual costs.

For DoD decision-makers, quantitative cost estimation is preferred as it often fosters a degree objectivity for the analysis. However, quantitative approaches are not always feasible. As noted by Mun (2016), "Qualitative forecasting is used when little to no reliable historical, contemporaneous, or comparable data exist" (p. 1). Mun (2016b) continues, "several qualitative methods exist such as the Delphi or expert opinion approach (a consensus-building forecast by field experts, marketing experts, or internal staff members), management assumptions (target growth rates set by senior management), as well as market research or external data or polling and surveys (data obtained through third-party sources, industry and sector indexes, or active market research)" (p. 1). Because these cost estimates are either single-point or are range estimates, there is a methodology that can augment qualitative data with more rigorous empirical cost estimating. For example, the prediction values from these qualitative methodologies can be inputted into cost estimation software in order to conduct parametric or non-parametric simulation that can augment risk assessment. It follows that this approach can “leverage experts’ knowledge by combining it with available quantitative data to arrive at more reliable costs estimate such as in ship-building cost estimates. Expert knowledge can be leveraged using the software by including qualitative estimates with quantitative analysis techniques” (Mun, 2016b, p. 1). This risk-based approach to cost is useful as it allows robust modeling to assess cost as well as schedule risk.

Integrated Risk Management Methodology

IRM is a system developed by Dr. Jonathan Mun designed to provide management the ability to analyze risk associated with the development of a new project or initiative. IRM combines several commonly accepted analytical procedures, such as predictive modeling, Monte Carlo simulation, real options analysis, and portfolio optimization, into a single, comprehensive methodology. The methodology uses existing techniques and metrics such as discounted cash flow, return on investment (ROI), and other metrics within the analytical processes to improve the traditional manner of evaluating potential projects within a company or the DoD. In contrast to the other methodologies, IRM focuses on the risk involved with a decision. It seeks to mitigate negative effects from risk while maximizing rewards from potential outcomes. At its core, IRM is a technique to provide managers the best analytic information available to use during the real options process.

There are eight steps within the IRM methodology:

1. Qualitative management screening
2. Forecast predictive modeling
3. Base case static modeling
4. Monte Carlo risk simulation
5. Real options problem framing
6. Real options valuation and modeling
7. Portfolio and resource optimization
8. Reporting, presenting, and updating analysis

While each of the individual steps provides value to a project manager, incorporating all of them in a contiguous approach will allow decision-makers the most effective use of the IRM process.

Figure 11 illustrates the comprehensive IRM process. The process begins with a qualitative management screening of potential projects, assets, and initiatives that could benefit the organization. These potential additions to a company's portfolio should align with the overall strategy, mission, and goals of the company (Mun, 2016a). The risks to an organization must be identified and addressed for decision-makers to have a realistic picture of the challenges the projects may face (Mun, 2016). This step is not unique to IRM. Prior to a firm beginning any venture, senior leadership should ensure that the ventures they are funding are realistic options based on their expertise and vision. If these are not in alignment, the initiatives will almost certainly fail. However, by evaluating the suitability of the projects and programs at the outset, management can eliminate potential programs that are incompatible prior to additional costly analysis.

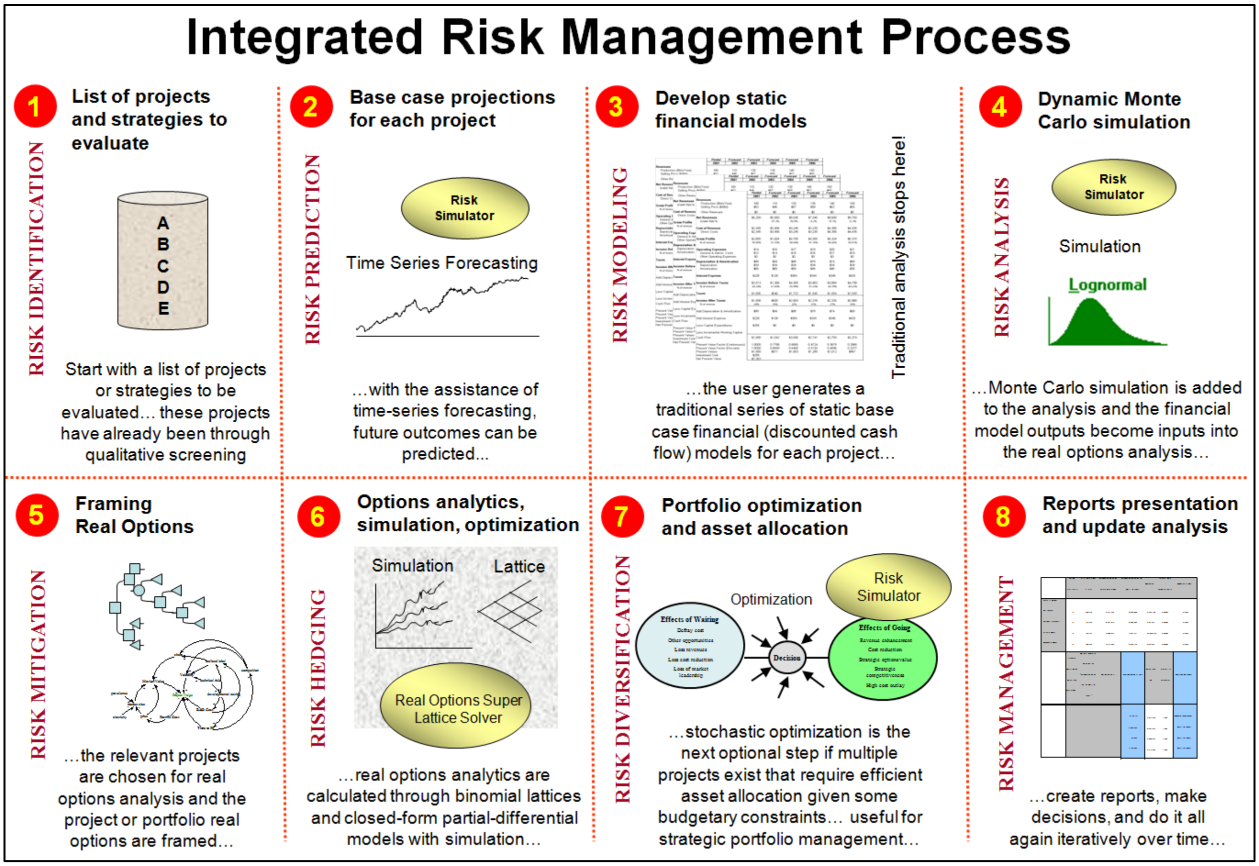


Figure 11. IRM process

Source: **Mun & Housel (2010)**

The second step is to forecast results using predictive modeling. Ideally, management will have access to historical data to use during this evaluation. Using comparable data from similar firms or projects is an acceptable alternative when the historical information is not available. With the data, analysts will use techniques such as multivariate regression analysis, time-series analysis, and others to predict a project's performance (Mun, 2016). If the data are unavailable, qualitative forecasting methods and SME estimates can be substituted for the historical or comparable information (Mun, 2016). The qualitative techniques can vary from assumptions about the growth rate to expert opinions, subjective estimates, and the Delphi method (Mun, 2016). In both cases, the techniques are forecasting value and cost drivers within the project (e.g., quantity, volume, production, revenue, cost, schedule; Mun, 2016). In a nonprofit context such as the DoD acquisition life cycle, surrogates should be used for revenue. The metrics that will define the value of a project can be projected in this analysis in place of for-profit financial measurements.

Using the results from the forecasting step, a model of discounted cash flow or similar models with a future projection of cost and benefit is created for each project, which serves as the base case analysis for future decisions (Mun, 2016). The net present value (NPV) or other ROI for the initiative is calculated via the traditional method, that is, projecting both revenue and cost and discounting the net value at an appropriate rate adjusted for standard financial risks (Mun, 2016). Additional profitability, productivity, and cost–benefit metrics, such as other variations of ROI, are calculated during this phase (Mun, 2016). The DoD and other nonprofit organizations do not collect revenue, making the profitability ratios listed meaningless without a surrogate for revenue. KVA offers this surrogate in the form of value. Using KVA as the base case analysis allows a quantitative, common-units comparison of nonprofit projects in the same manner as a traditional revenue-generating industry.

Next, the analyst will conduct a Monte Carlo risk simulation to obtain a better assessment of the potential risks and value of the proposed venture. While the base case static model developed in step three is a useful tool, it is based on static information and, as such, produces a single-point estimate (Mun, 2016). The information gleaned from the model may not be accurate due to the uncertainty and risks involved in future cash flows (Mun, 2016). Since financial problems inherently contain uncertainty of some form, a model that accounts for this uncertainty is necessary (Brandimarte, 2014). The Monte Carlo simulation will increase confidence in the value of a project by using statistical analysis to give a probability of ranges for different variables.

Monte Carlo simulation, or probability simulation, is a technique used to understand the impact of risk and uncertainty in financial, project management, cost, and other forecasting models (Mun, 2016). In a Monte Carlo simulation, analysts generate random scenarios and gather relevant statistics to assess situations that are affected by uncertainty (Brandimarte, 2014). Using historical data and the opinions of SMEs, analysts can input a range of possible values to simulate potential future outcomes (Mun, 2016). Since the input variables are given in a range of estimates, the model's outputs will also be a range indicating the likelihood of the possibilities (Mun, 2016). The Monte Carlo simulation can also be run using only historical data, and the computer will make a custom distribution of the variables to produce its output or with a prescribed probability distribution (Mun, 2015). In IRM, the analyst will set NPV or any of the computed ROI variations as the resulting variable(s) and run the Monte Carlo simulation thousands of times, adjusting each of the other variables to predict a range and probability of potential NPVs for the project (Mun, 2015).

The quantitative data gleaned from the Monte Carlo simulation is only useful if it provides decision-makers with improved information to make decisions. The information must be converted into actionable intelligence (Mun, 2016). While the statistical analysis and other preceding steps are important, the crux of the IRM methodology is the real options assessment. To begin that process, leaders must conduct real options problem framing, step five in the IRM methodology. Real options allow managers to hedge, value, and take advantage of risks, reducing the potential downside while maximizing potential gains from volatile projects (Mun, 2016). By framing the problem through a real options lens, an organization's leadership can generate a strategic plan for the problem from several options (Mun, 2016). Analysts will then examine chosen options in more detail (Mun, 2016).

Real options provide investors the ability to adjust the course of previous decisions based on the performance of the investment to date. They allow management to make "better and more informed strategic decisions when some levels of uncertainty are resolved through the passage of time, actions, and events" (Mun, 2015, p. 438). Options are opportunities for a company; they have a right to conduct an action without the obligation to take the future action (Dixit & Pindyck, 1995). There are several types of options, and the number of names of available options varies depending on the literature source. Some of the more common categories are briefly covered below.

The *option to delay* gives managers the ability to adjust the timing of a project (Damodaran, 2000). When analyzing the cash flows of a project, a negative NPV or ROI indicates a project is not a good investment at the current time (Damodaran, 2000). As illustrated in Figure 12, waiting until the NPV turns positive allows an organization the option to delay the initiative until it will benefit the company. (NPV is not the sole source to make an option decision within IRM and is included to illustrate the concept in a simple manner.) The statistical analysis conducted in previous steps allows analysts to determine the optimal time to make project investment decisions. This option is also referred to as a deferment option, option to wait, or option to execute (Mun, 2015). The option to delay is often executed through pre-negotiated prices or similar contracted terms that offer the choice to purchase something without an obligation to do so (Mun, 2015). These terms could include options based on a build, buy, or lease contract; a proof of concept test; market research; research and development; or other negotiated terms (Mun, 2015).

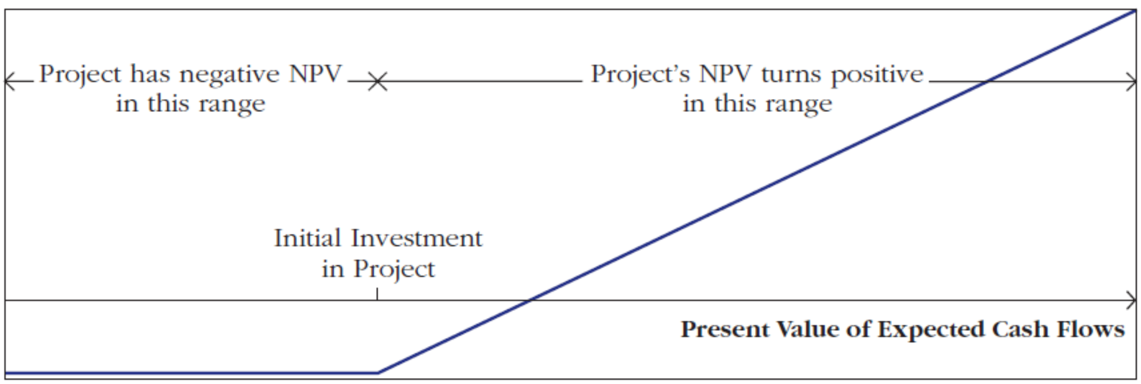


Figure 12. The option to delay

***Source:* Damodaran (2000)**

The *option to abandon* a project provides management a way to reduce future losses in a project that is not performing as anticipated (Damodaran, 2000). Figure 13 shows one example when the option to abandon should be considered. As the present value of the project decreases below the liquidation or salvage value of the project, managers should abandon the project and salvage as much as possible from the existing infrastructure and investment (Damodaran, 2000). Salvage is not the only way to execute the option to abandon. Companies can also execute the option to abandon through contractual buyback provisions, termination for convenience, divestitures, or early exit clauses (Mun, 2016).

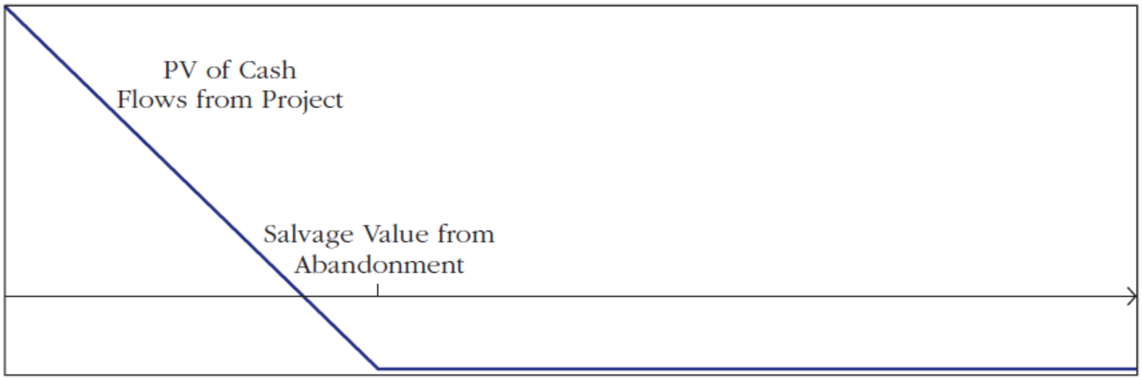


Figure 13. The option to abandon
***Source:* Damodaran (2000)**

A third real option available to leaders is the *option to expand* (Damodaran, 2000). In this instance, an investment in a project allows a company to undertake additional projects or to enter new markets, expanding the scope of the original investment (Damodaran, 2000). While not always the case, businesses may be willing to accept a negative NPV for the initial project to have access to the expansion options it will create with the promise of higher NPVs (Damodaran, 2000). By investing in the original initiative and maintaining the option to expand, the company is limiting the potential upside from an initial investment into the entire project; however, it is also reducing the downside risk of a failed, high-capital investment (Damodaran, 2000). For example, a company may recognize a potential market in creating a suite of sensors for a new autonomous vehicle. Without the existing infrastructure to compete in this market, leadership decides to develop a project that will create a single sensor for the vehicle. When the project is completed, managers assess the financial feasibility of creating additional sensors. The original investment must be a requirement for the subsequent project to be an option to expand. That is, the additional sensors could not be developed without the investment into the first sensor. Otherwise these are simply a collection of separate but related projects.

Other real option strategies include barrier options, chooser options, contraction options, sequential options, and switching options (Mun, 2016). *Barrier options* become available when an artificial barrier is either breached or not breached (e.g., profits exceed a certain level or vendor prices fall below a specified threshold; Mun, 2015). *Chooser* *options* permit management to choose between one or multiple strategies, such as expanding, abandoning, and so on (Mun, 2015). *Contraction options* allow a firm to contract its existing operations to cut operating expenses under certain conditions (Mun, 2015). This could happen through outsourcing, subcontracting, leasing, or other alternatives (Mun, 2015). *Sequential options* require a previous option to successfully finish prior to initiating a subsequent option, compounding the options, and reducing the downside risk from a large upfront investment (Mun, 2015). Finally, *switching options* provide management the ability to switch operating conditions, such as technologies, markets, or products (Mun, 2015). This type of option gives a firm strategic flexibility in choosing a course of action, keeping its current project while exploring possible substitutions (Mun, 2015).

After determining which real option may be appropriate, analysts conduct simulations on the chosen options to complete the real options valuation and modeling. The results from the Monte Carlo simulation and previous evaluations give a probability distribution of values that illustrate the uncertainties and risks associated with each project, which, when combined, give a distribution of the NPVs and the initiative's volatility (Mun, 2016). The assumption within a real options context is that future profitability of the project is the fundamental variable of interest, measured by future cash flow series (Mun, 2016). Analysts use the future cash flow and the present value of the future cash flows to determine the total asset value of the project in a real options model (Mun, 2016).

The real options analysis reveals the financial and economic strengths and weaknesses of the project's available strategic options, allowing analysts to make recommendations to management on which projects to pursue. Projects are typically not conducted individually within businesses, and initiatives are often correlated (Mun, 2016). If managers view the future projects as a portfolio, they can hedge and diversify the risks associated with each singular project (Mun, 2016). Using traditional portfolio analysis will assist leadership in determining the optimal allocation of investments throughout their collection of projects (Mun, 2016).

Generating coherent and concise reports detailing the analysis is the eighth, and final, step in IRM (Mun, 2016). If decision-makers do not understand the complicated procedures that led to the investment recommendations, they will not trust the results enough to follow those recommendations (Mun, 2016). Transforming the "black-box set of analytics into transparent steps" is vital to ensuring leadership has the best possible information with which to make decisions for the company's project portfolio (Mun, 2016, p. 95). Although this is the final step within the IRM process, as additional information becomes available and the uncertainty and risk are reduced or resolved, analysts should revisit the models with updated information (Mun, 2016). Reworking the original models with the new data allows managers to make midcourse corrections to improve the performance of both the individual project and the portfolio of projects (Mun, 2016).

The IRM methodology is a systematic technique to determine the best possible projects to pursue based on the statistical likelihood of their success. Using historical knowledge of defense acquisition programs and AI systems in both the government and commercial realms could improve the budgeting and scheduling processes. Determining the likely range of outcomes through dynamic statistical modeling may improve the program's performance. By better understanding the risk associated with various components, a more appropriate schedule and budget could be developed. IRM may also help determine which real options should be included in acquisition contracts. A high-risk program may need more options, such as the options to abandon, delay, or expand, based on its actual performance. Finally, IRM could prove useful in portfolio management, helping decision-makers determine which programs to initiate when viewing the portfolio of other programs in progress and used operationally.

THIS PAGE LEFT INTENTIONALLY BLANK

Artificial Intelligence

The purpose of this section is to discuss Artificial intelligence (AI) in context of its current impact in the Defense Acquisition System. This is achieved by providing a brief history of AI, how it relates to the international security environment, and then relating this information to the previous section by discussing the challenge that AI technologies bring to the DoD due to its complexity. Finally, the section ends with a discussion on how the three Acquisition methodologies (EVM, KVA, and IRM) can be utilized to assess an AI program based on its stage in the Acquisition life cycle.

Artificial Intelligence Defined

AI does not refer to a specific system. It is a broad nomenclature for a collection of related inorganic computer science methods used to simulate human intelligence. The term *AI* typically conjures up the general concept of machine learning, which, in reality, is a type of AI where a computer system is programmed to identify and categorize external real-world stimuli via a “learning” process. The DoD’s AI strategy defines AI as “the ability of machines to perform tasks that normally require human intelligence—for example, recognizing patterns, learning from experience, drawing conclusions, making predictions, or taking action—whether digitally or as the smart software behind autonomous physical systems” (DoD, 2019b). This capability of enhanced automation is of great interest to the DoD as potential future near-peer adversaries such as Russia and China are investing heavily in this field for military purposes (DoD, 2019).

The current research study focuses on acquisition of AI capabilities that are deemed most appropriate for the acquisitions context, which may include applications of Machine Learning (ML)—both supervised and unsupervised—and less volatile applications as Natural Language Processing (NLP), and more robust expert systems such as Robotic Process Automation (RPA), shown in Figure 14. The figure shows that AI is a combination of AI sciences, such as ML and NLP, and RPA, which benefits from the use of AI capabilities. (RPA is not simulation of human intelligence, but rather it just mimics capabilities).

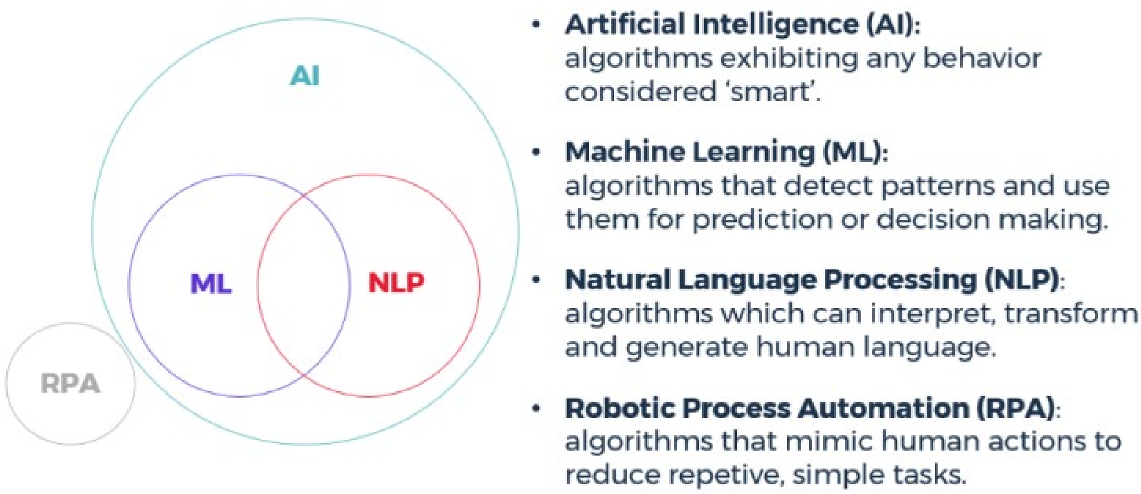


Figure 14. AI terms and relationships
Source: **Sievo (2019)**

For a more thorough discussion on the AI capabilities of ML (and its subsets), NLP, and RPA, see Appendix A. Payne (2018) defines these current capabilities that use ML, NLP, and RPA as "tactical AI" since these AI systems are "most applicable in the military domain to problems at that level of warfare—including those of maneuver and the application of fires" that "demand cognitive abilities of the sort that modern AI is already pretty good at: pattern-recognition, probabilistic reasoning, memory, and, above all, speed" (p. 164).

Overview and Brief History of AI

The terms *artificial intelligence* and AI are often misunderstood. This is due to the fact that the term is often confused with its three evolutionary stages, Artificial Narrow Intelligence, Artificial General Intelligence, and Super Intelligence (Kaplan & Haenlein, 2019; Yao et al., 2018). Artificial General Intelligence and Super Intelligence are still in the realm of science fiction where the machine has the capability of meeting or surpassing the cognitive abilities of the human brain (Bostron, 2014). For example, Kurzweil (2000) predicted two decades ago that mankind is approaching the era where computer intelligence will eventually surpass human cognition. However, in the literature, the basic definition of artificial intelligence imputes capabilities that are characterized by the General or Super Intelligence level as in small world contexts; "Artificial Intelligence is intelligence exhibited by machines in a given task that is similar to or the same kind as human intelligence in the same task" (Li et al., 2019, p. 3619). While the definition notes the specificity of the task, comparable to human intelligence required to perform the task, this level of intelligent behavior is not characteristic of most forms of AI.

The definition of AI that is adopted for the current research study, falls within the stage Narrow Intelligence: "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (Haenlein & Kaplan, 2019, p. 5). This definition traces its roots to the mid-20th century, primarily to Norbert Weiner's (1948) concept of "Cybernetics" which he described as teaming between the human and machine. In order to defeat the threat of the German supersonic V-2 Rockets during the Battle of Britain in World War II (WWII), Weiner and his colleagues proposed a systemic framework of unifying a system between the human analysts and the machine, their radar equipment and anti-aircraft guns (Capra, 1996). William Ashby (1956) expanded upon the concept of cybernetics to discuss how machines could amplify human intelligence in what he described as "intelligence amplification." This was followed by the development of the concepts of Licklider's (1960) conceptualization of "Man-Computer Symbiosis" and Englebart's (1962) "Augmented Human Intellect."

It was during this period that the term *Artificial Intelligence* started to diverge from the larger field of cybernetics. The subject of AI became its own field of study that was formally founded at Dartmouth in 1956. The science of AI was created to determine if inorganic machines could perform human-level intelligent functions (Denning, 2019; Heller, 2019). It went through several hype cycles, due primarily to sensationalizing of what it might be able to do, with frequent disappointments (Figure 15). Significant enthusiasm for AI reemerged at the same time that Big Data computing power became more accessible to researchers and companies, which, in turn, could apply the AI science to multiple tangible applications (Haenlein & Kaplan, 2019). Currently, examples of commercially viable applications of AI exist in manufacturing robots, smart assistants, proactive healthcare management, disease mapping, automated financial investing, virtual travel booking agents, social media monitoring, conversational marketing bots, NLP tools, and contract management (Daley, 2019).

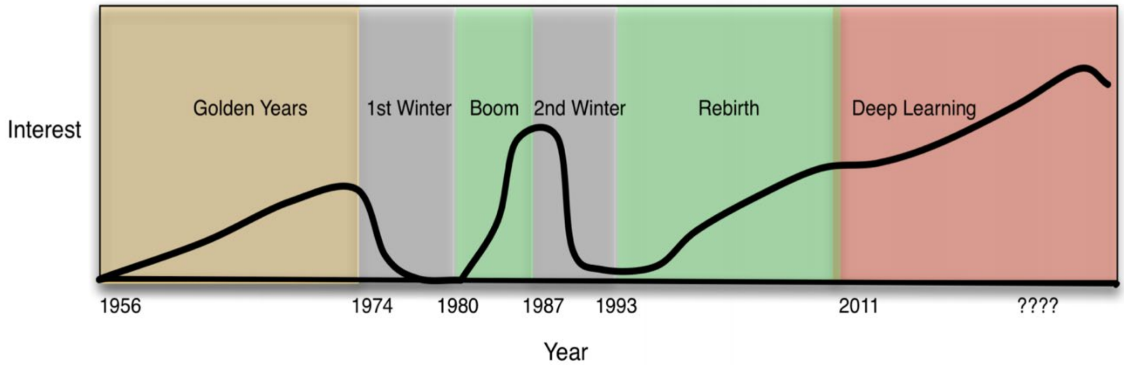


Figure 15. The timeline of interest in AI during different phases of its development
Source: **Denning (2019)**

As Figure 15 depicts, AI was a novel discussion during the post-WWII era when it diverged from the field of cybernetics. However, the interest in AI went through two cyclical interests drops in the 1970s and again in the 1980s, thereby leaving the AI discussion primarily to the academics and science fiction authors. Haenlein and Kaplan (2019) explain the reason for these hype cycles. Since its initial establishment, "AI remained an area of relative scientific obscurity and limited practical interest for over half a century. Today, due to the rise of Big Data and improvements in computing power, it has entered the business environment and public conversation (Haenlein & Kaplan, 2019, p. 5). Despite being conceptually viable for academia in the "Golden Years" of the 1950s, it was not until the second decade in the 21st century that the computing technology and Big Data sets were available for industry to realize the possibilities that were proposed in the early days of AI (Denning, 2019; Yao et al., 2018). Thus, society has entered into an era where Deep Learning made possible through AI can expand the capabilities of the military in data analytics, image recognition, and human machine teaming, and enhance cyber warfare in both defensive and offensive operations (Denning, 2019).

The Growth of AI Literature from Inception to Industry 4.0

Utilizing the Web of Science (n.d.) comprehensive academic search engine, the researchers found 316,009 scholarly publications, including 188,275 academic journal articles on the topic of AI. The period of these AI publications covers the entire timeline from the inception of term *artificial intelligence* in the early 1960s until present day. Using

Van Eck and Waltman’s (2020) VOSviewer tool for visualizing scientific landscapes, the researchers created a network map that illustrates how AI (“intelligence” in the diagram) relates to other key topics. A visualization of the major key terms in AI research is depicted below in Figure 16.

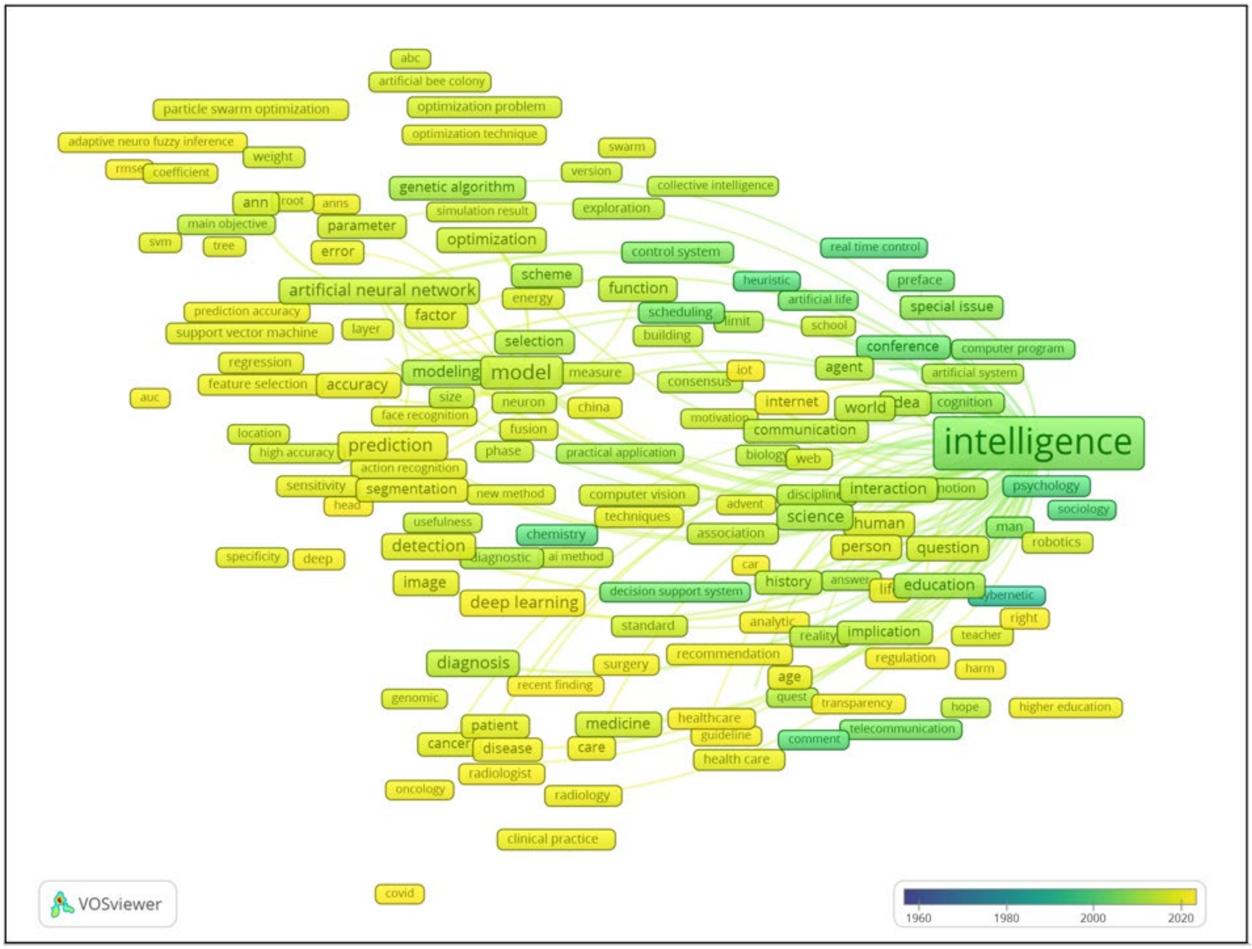


Figure 16. The Key Topics in AI from 1960 to Present Day

Figure 16 depicts how AI relates to numerous terms that have more than 50 publications related to AI since it branched out of cybernetics in the 1960s. Figure 17 provides a clearer portrayal of the terminology using network nodes.

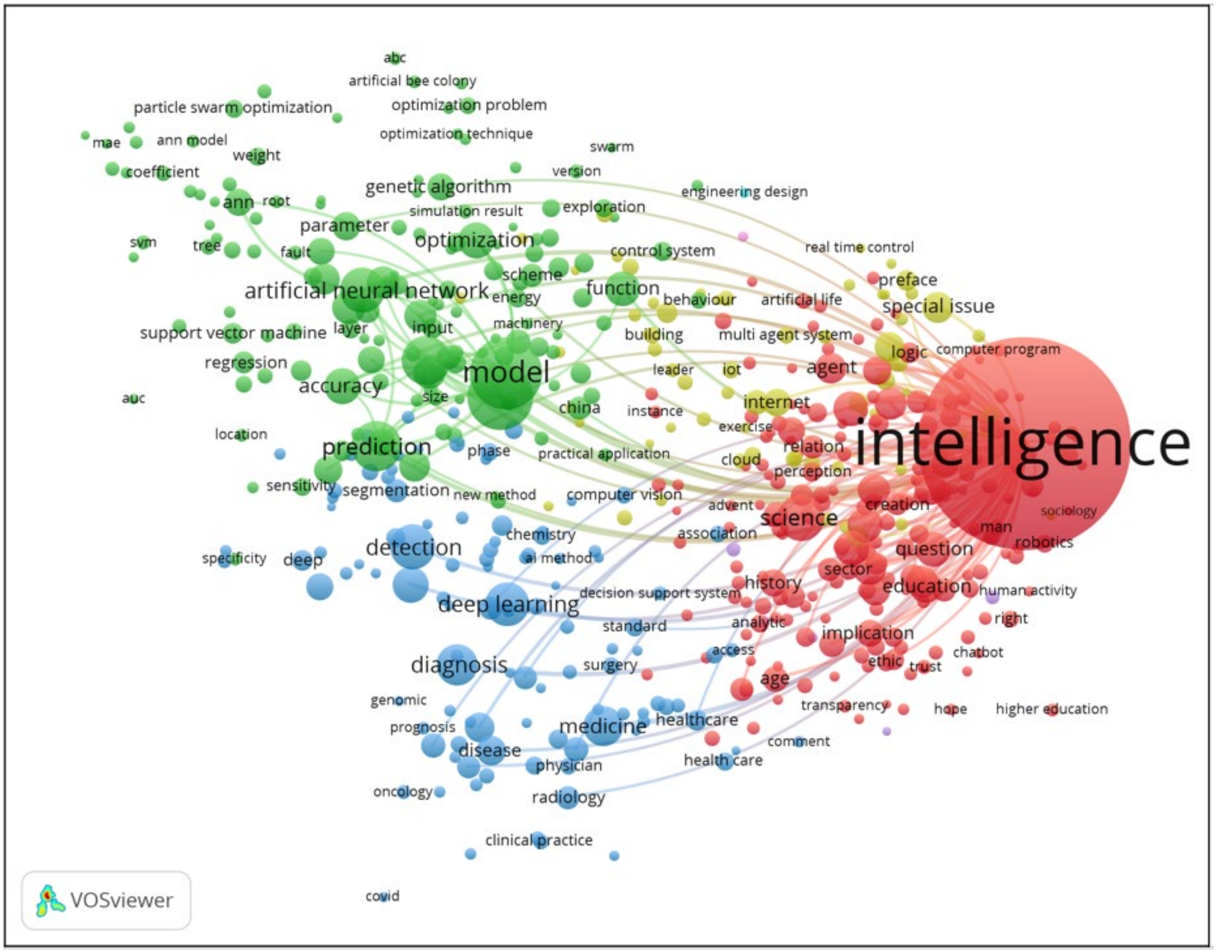


Figure 17. The Network Typology of AI Topics

As Figure 17 depicts, AI topics have branched out of the general computer science and automation fields into other fields such as healthcare, as noted in medical terms in Figure 17. As the timeline illustrated in Figure 16, much of the expansion in AI topics has occurred in the past two decades.

The expansion of AI into new fields, as illustrated in Figures 16 and 17, is shown in the growth of research areas. The bar graph in Figure 18 shows the number of scholarly publications on the topic of AI by the top 10 research areas. Not surprisingly, the majority of the AI publications are in the research area of "Computer Science" with more than 286,000 publications. Equally, of no alarm is the fact that nine of the top 10 research areas for AI are in the science, technology, engineering, and mathematics (STEM) fields (see Figure 18).

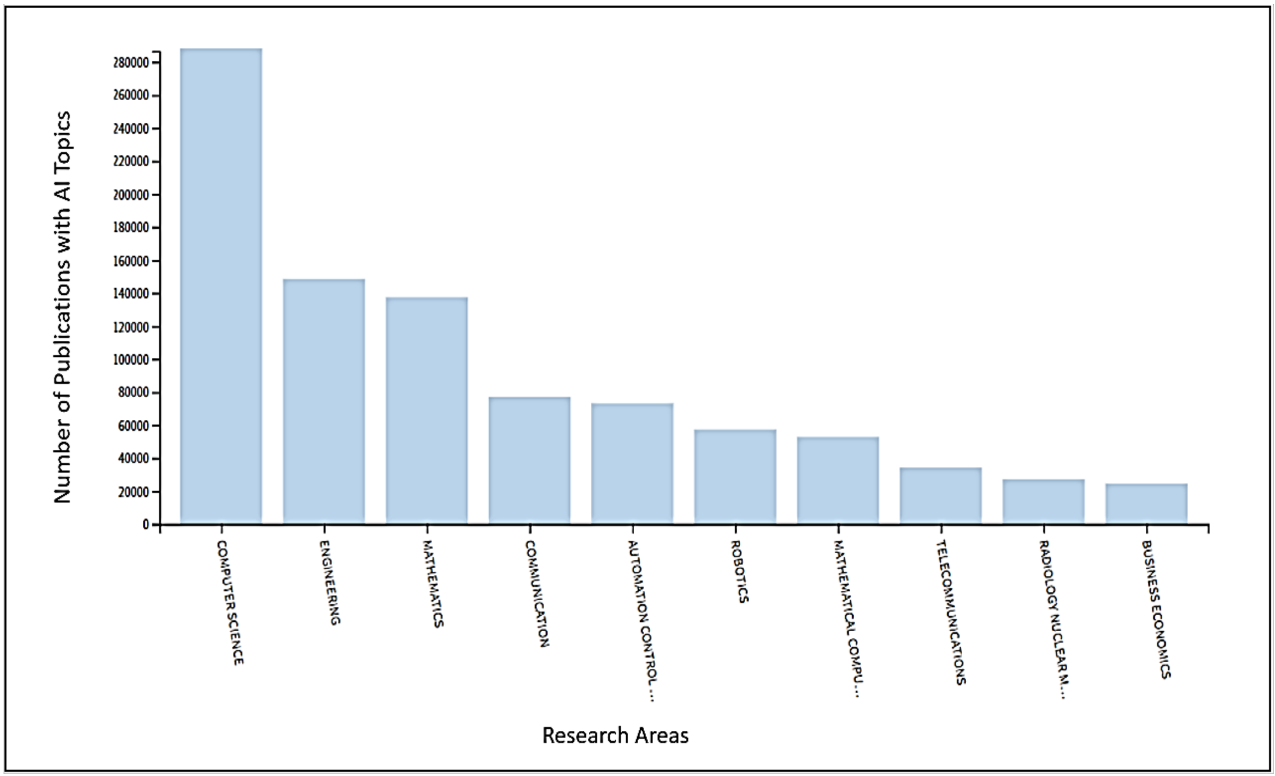


Figure 18. Bar Graph of Number of AI Publications by Research Areas

Despite being the smallest research area, the growth in AI publications in the Business and Economics area in the past two decades is perhaps the most revealing although the nearly 23,000 AI publications in the Business and Economics area is a small fraction of more than 300,000 AI publications since 1960. However, this is the research area that has seen the largest growth in the past two decades. Burkhalter (1963) wrote the first AI business and economics article on applying AI for pattern recognition problems in operations research. Burkhalter (1963) was only article that discussed the use of AI in the business and economic fields. Much of the next four decades was an incremental expansion of AI based off the cycle in interest in AI. Thus, the business scholarly publications in AI next exceeded double digits until 1990 when the publications jumped from 74 in 1989 to 126 in 1990. This growth fell in line with the "Rebirth" of interest in AI as depicted in Figure 15 (Denning, 2019). This growth in AI research in the business field has expanded since then, particularly in the past decade to the point that in 2019, there were nearly 3,000 business and economics publications in the topic of AI.

A major cause of the business interest in AI in the past decade is the phenomenon known as “The Fourth Industrial Revolution” or “Industry 4.0” (Lee, 2013; Schwab, 2015). In noting how steam power impacted the agrarian society in the First Industrial Revolution, the assembly line in the Second Industrial Revolution, the computers in the Third Industrial Revolution, Schwab (2015) argues that society is on the cusp of a “technological revolution that will fundamentally alter the way we live, work, and relate to one another.” Schwab (2017) and his colleagues at the 2011 Hanover Conference on Technology made the case that impact of this new revolution is quite extensive:

> We have yet to grasp fully the speed and breadth of this new revolution. Consider the unlimited possibilities of having billions of people connected by mobile devices, giving rise to unprecedented processing power, storage capabilities and knowledge access. Or think about the staggering confluence of emerging technology breakthroughs, covering wide-ranging fields such as artificial intelligence (AI), robotics, the internet of things (IoT), autonomous vehicles, 3D printing, nanotechnology, biotechnology, materials science, energy storage and quantum computing, to name a few. Many of these innovations are in their infancy, but they are already reaching an inflection point in their development as they build on and amplify each other in a fusion of technologies across the physical, digital, and biological worlds. (Schwab, 2017, p. 9)

Kovacs and Kot (2016) highlight that the “essence of Industry 4.0 conception is the introduction of network-linked intelligent systems, which realize self-regulating production: people, machines, equipment and products will communicate to one another” (p. 122). As society continues the journey into The Fourth Industrial Revolution, it can be assumed that fields outside of STEM will also continue to research and explore more uses for artificial intelligence. Thus, while much of this research focuses on the DoD Acquisition of AI, the tools this research proposes for the project management in developing and procuring AI (KVA and IRM) have a broader application.

AI in National Security and Defense Applications

As the capabilities of AI has expanded in the Fourth Industrial Revolution, there has been a growing concern that the international arena, particularly the three Great Powers of the United States, China, and Russia may already be in the throes of an “AI Arms Race” (Geist, 2016). One of prevalent fears among scholars in the age of expanding

AI is the risk of developing “autonomous weapons” that can no longer be controlled (Geist, 2016). In essence, by developing AI weapon systems, humanity may be sowing the seeds of its own destruction. As former Secretary of State Kissinger noted on the rise of AI:

> The scientific world is impelled to explore the technical possibilities of its achievements, and the technological world is preoccupied with commercial vistas of fabulous scale. The incentive of both these worlds is to push the limits of discoveries rather than to comprehend them. And governance, insofar as it deals with the subject, is more likely to investigate AI’s applications for security and intelligence than to explore the transformation of the human condition that it has begun to produce. (Kissinger, 2019)

Kissinger’s (2019) words strike at the heart of the growing mistrust of this new technology that has exploded with the recent rise of Big Data and more powerful computing. This fear may also grow as AI can potentially dominate international relations with a new race to develop and weaponize AI (Geist, 2016).

Technology Trust and AI

As AI is an emerging technology with unlimited possibilities that have only been conceptualized in the realm of science fiction, a major consideration with AI is the issue of Technology Trust. In a speech to Russian students in 2017, Russian President Vladimir Putin famously stated, “Artificial intelligence is the future, not only for Russia, but for all humankind. It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world” (Dougherty & Jay, 2018, p. 1). This statement articulates the general belief among policy experts and decision-makers that AI can transform and revolutionize the international security environment in a manner unlike any emerging technology since the introduction of the atomic bomb. The emerging concept of concerns for Technology Trust in regard to AI came to light in 2007 when a British technology professor wrote, “We are sleepwalking into a brave new world where robots decide who, where and when to kill” (Sharkey, 2007). As a response to the evolving capabilities in unmanned and autonomous systems, scholars in the technology field pushed for a ban on AI on the battlefield. The most prominent ban was in April 2013 when an activist movement called the “Campaign to Stop Killer Robots” began to gain traction in international forums (Gubrud, 2014). While many laughed off the concept of “Killer Robots” as something limited to the realm of science fiction created by Hollywood, some of the world’s greatest modern thinkers, such as the late Stephen Hawking and Elon Musk have been very outspoken against the risks to AI (Burton & Soare, 2019; Gibbs, 2014). Many of these thinkers signed an open letter in 2015 that described the existential threat that AI places on humanity (Sparkes, 2015). Rather than continuing to develop weaponized AI, the open letter proposed a push for “expanded research aimed at ensuring that increasingly capable artificial intelligence systems are robust and beneficial: our artificial intelligence systems must do what we want them to do … such research directions that can help maximize the societal benefit of artificial intelligence” (Future of Life Institute, n.d.). The similarities of the proposed banning of weaponized AI harkens back to the Cold War era and the social and political movements that pushed for a similar banning of nuclear arms. This is because AI, with the possibility of developing autonomous weapon systems, has the potential to dominate the strategic landscape unlike any emerging technology since the atomic bomb (Geist, 2016).

What makes AI so much more novel than previous technological developments is an inherent risk that the technology is able to become autonomous. As noted by Mun and Anderson (2020), “There are many issues associated with trust in technology that are increasing in importance as the U.S. military begins to acquire and deploy autonomous systems” (p. 3). The authors note that the challenge of the AI is that as its technology advances and moves toward capability to deploy autonomous systems, the greater the risk from DoD autonomous systems, particularly weapon systems. Therefore, the authors propose that “there is a need to establish a system of metrics that justify a level of technology trust” (p. 3). A key factor in ensuring that the DoD can gain trust in novel technologies, such as AI, over time is to continue research into the field. As AI develops in capability from the realm of science fiction and into the realm of science, it becomes increasingly more important to establish Technology Trust as early as possible.

Related to Technology Trust is the ever-present concern that AI would develop to the point that it met the threshold of human intelligence (Yao et al., 2018). Thus in 1950, Alan Turing developed a test to assess the capabilities of AI: “If a human is interacting with another human and a machine and unable to distinguish the machine from the human, then the machine is said to be intelligent” (Haenlein & Kaplan, 2019, p. 7). Thus, the Turing Test is a benchmark assessment often used in AI to see if a computer program can fool people into thinking it is human (Oppy & Dowe, 2016). For a more detailed discussion on Technology Trust, including the Defense Advanced Research Projects Agency's (DARPA) role in developing explainable reasoning within the algorithms, see Appendix A.

Despite the numerous risks associated in developing and fielding AI systems in the battlefield, the nature of the global competition and AI arms race with China and Russia require the DoD make the strides to develop AI capabilities. The arena that shapes international security necessitates a risk averse culture, particularly in dealing with peer or near-peer competitors such as Russia and China. As once stated by the late Reinhard Selten, 1994 Laureate in the Nobel Memorial Prize in Economics, "Generals cannot make the *one* mistake" (personal communication, 2006, emphasis supplied). This is likely why Selten (1991) stated that the "advantage of modelling conflicting views about international policy as games with incomplete information is that it helps to yield very new explanations of why different countries choose different policies or why different parties of the same country follow different strategies" (p. 253).

The implication in Selten's statement and his work in Game Theory is that the risks of failure that military leaders face are so great that caution is the norm in the non-cooperative game theory in the national security. Such was the case in 1983 war scare where the Soviet military commanders mistakenly saw the Able Archer 83 NATO exercise as mobilization toward war with the Warsaw Pact (Able Archer 83 Sourcebook, 2019). If it were not for U.S. Air Force Lieutenant General Leonard Perroots recognizing the Soviet Union's maneuvers as leading toward hostility and thereby de-escalating NATO forces, the "mistake" would have been quite severe (Roberts, 2017).

Thus, the implications in applying non-cooperative game theory as in the use case of Able Archer 83 during the Cold War is that the stakes involved in the international security arena are tremendously high. Therefore, military leaders and DoD decision-makers should be willing to make risks in developing systems that use highly volatile technologies such as AI. Therefore, the White House's (2020) recently published *National Strategy for Critical and Emerging Technologies* states the following:

> Some emerging technologies are globally diffuse or are too early in the R&D phase to have clearly identified implications for United States national security. In those cases, a risk management approach will be applied to gauge national security implications, inform investments, and monitor development. In managing risk, the United States Government will first identify, evaluate, and prioritize its technology risks, followed by a coordinated response to avoid, reduce, accept, or transfer risk. (White House, 2020, p. 4)

In the *National Strategy for Critical and Emerging Technologies*' Appendix, one of 20 key technologies that is highlighted as a critical, emerging technological area is AI. Therefore, the realities of the international arms race necessitate that the DoD make advances toward developing its AI capabilities. Thus, the volatilities of creating AI systems will require risk management tools such as those proposed in this research (i.e., IRM).

AI and the DoD's Third Offset Strategy

Now that this paper has provided an overview of AI, particularly in relations to the national security along with the prevalent risk of autonomous weapon systems and inherent mistrust of this emerging technology, the discussion will now shift toward the DoD's recent drive toward AI. As the potential capabilities of AI became more evident, then-Secretary of Defense Chuck Hagel (2014) launched the Third Offset Strategy, which is described as follows:

> An ambitious department-wide effort to identify and invest in innovative ways to sustain and advance America's military dominance for the 21st century. It will put new resources behind innovation, but also account for today's fiscal realities—by focusing on investments that will sharpen our military edge even as we contend with fewer resources. Continued fiscal pressure will likely limit our military's ability to respond to long-term challenges by increasing the size of our force or simply outspending potential adversaries on current systems, so to overcome challenges to our military superiority, we must change the way we innovate, operate, and do business. (Hagel, 2014b)

The DoD's Third Offset Strategy hearkens back to the First Offset launched by President Eisenhower where the United States developed its strategic arms to reduce the need for standing conventional forces during the early Cold War and the Second Offset where after the Vietnam War, the DoD, through programs such as DARPA, increased the capabilities of its conventional forces to counter the Warsaw Pact (Hillner, 2019). Then-

Deputy Secretary Work noted that it is in the challenges of the current geopolitical environment that artificial intelligence can help reinstate the United States' previous technological overmatch, "Learning machines are an example of technology that can help turn AI and autonomy into an offset advantage" (Pellerin, 2015). In a separate speech, Work (2014) stated that it was the Third Offset Strategy that pushed for subsets initiatives such as the Defense Innovation Initiative and then-Under Secretary of Defense for Acquisition, Technology and Logistics Frank Kendall's Better Buying Power 3.0, which were focused on bringing back the competitive edge that the U.S. military once had over its geopolitical competitors. Voelz (2016) notes that it is through the Defense Innovation Initiative and Better Buying Power 3.0 that "focuses on achieving high-payoff breakthroughs in areas such as artificial intelligence, robotics, additive manufacturing, and nanotechnology" as well as the establishment of the "Defense Innovation Unit Experimental in Silicon Valley to 'scout, connect, and support the innovation of disruptive technology' with potential military value" (p. 180). Voelz (2016) states that a "common theme among these initiatives" (p. 180).

The theme that arises from the Third Offset, Defense Innovation Initiative, and Better Buying Power 3.0 is that the Defense Acquisition System is unable to meet the requirements of fielding software intensive systems such as AI systems. Kendall (2017) discussed this problem by illustrating the challenges that PMs with software intensive systems such as AI face due to the extensive cycle of developing, testing, and fielding "several builds of software in various stages of maturity" simultaneously while dealing with the organizational bureaucracy that slows the PM down (p. 50). Thus, Kendall (2017) concludes that "in attempting to implement Agile software development practices this PM has run into constraints from MAIS [Major Automated Information System] and DoD acquisition processes that have stymied modern software development best practices. … We're getting in his way" (p. 51). In a similar tone, through a recent article in *War on the Rocks*, the chair and vice-chair of the National Security Commission on Artificial Intelligence solicited support from academics and military professionals on how to best implement AI. One of the key questions they asked is, "What acquisition and application processes need to change to allow the government to address the AI national security and defense needs of the United States?" (Work & Schmidt, 2019). Thus far, the answer from those fielding the systems is to rely heavily on prototyping rather than the traditional Acquisition process (DiNapoli, 2019). Furthermore, those in the field recommend taking advantage of the dual-use technologies in the private sector. As recommended by an article in an "AI and National Security" special series of articles in *War on the Rocks* that responded to the chair and vice-chair of the National Security Commission on Artificial Intelligence:

> The Department of Defense and intelligence community should continue to adapt their approach to better capture innovation happening in the commercial ecosystem. Initial steps by U.S. Special Operations Command to streamline procurement and contracting processes are a good start, such as the SOFWERX Data Engineering Laboratory. But deeper collaboration and teaming across the Department of Defense, private sector, and academia is required to develop the data culture and architecture necessary for success. (Egel et al., 2019)

DoD Acquisition & AI

DoD Acquisition of New Software-Intensive Technology

The fielding of new and advanced technologies such as AI is a challenge for the DoD and all federal government. The current methodologies have proven unsuccessful in meeting the task of providing the requirements to the warfighter to face the challenges of the modern battlefield (Kendall, 2017). However, with the release of recent strategic changes such as implementing innovation practices and advanced prototyping, the DoD may prove up to the task of fielding the materiel and equipment to support the Department and the Services (Kendall, 2017; Voelz, 2016). Some proven, successful practices are adopting best practices in DoD labs and increasing the use of rapid development cycles through prototyping (DiNapoli, 2019; Sullivan, 2018).

A recent case study on the two-decade process of developing biometrics for use in the Services provides the lessons learned for acquisition of new and advanced technology. According to Voelz (2016), "The case study of biometrics demonstrates that effective military innovation can only occur through an integrated approach that takes into account the interdependent elements of technology development, acquisition planning, doctrinal design, and warfighting strategy" (p. 180). Despite having no prior experience in biometrics, the military was able to implement and adapt biometrics successfully as noted:

> In terms of rapidly developing and fielding a new technology, the record of defense biometrics should be considered a tactical success. During the course of the conflicts in Iraq and Afghanistan, U.S. forces generally made effective use of an emerging capability that directly enabled new forms of identity-based operations in response to unique demands of waging irregular warfare. However, the rapid fielding process did reveal shortcomings in how DoD manages military innovation at the bureaucratic level. These challenges are undoubtedly not unique to biometrics and are certainly worthy of future study to better understand how DoD can improve process models for wartime innovation. (Voelz, 2016, p. 195)

Thus, while not exactly the same, the lessons learned of adapting biometrics in the Services is an example of how the DoD can adopt AI throughout the department because biometric systems and AI systems both require heavily software intensive development processes. For such systems, the Defense Science Board (2018) study on the DoD's Acquisition of software intensive tools recommended that the

> Department of Defense (DoD) and its defense industrial base partners need to adopt continuous iterative development best practices. The study recommends DoD adopt best practices on risk reduction and metrics in formal program acquisition strategies. Software strategies must be better incorporated in current and legacy programs from development, production, and sustainment. (Defense Science Board, 2018)

DoD Coordination of AI through the JAIC

As noted earlier in the report, one of the challenges of implementing AI is that it encompasses so many areas in augmented human intelligence, machine learning, deep learning, big data, and neural networks (Haenlein & Kaplan, 2019). In practice, creating an approach toward "AI" is akin to the Defense Acquisition System lumping all weapons that fire a projectile into a single acquisition methodology. Therefore, to meet this issue, the DoD made several recent strategic initiatives. These include

- Establishing a Joint Artificial Intelligence Center (JAIC), which will "coordinate the efforts of the Department to develop, mature, and transition artificial intelligence technologies into operational use."
- Publishing a strategic roadmap for AI development and fielding, as well as guidance on "appropriate ethical, legal, and other policies for the Department governing the development and use of artificial intelligence enabled systems and technologies in operational situations."

- Establishing a National Security Commission on Artificial Intelligence to conduct a comprehensive assessment of militarily relevant AI technologies and provide recommendations for strengthening U.S. competitiveness. (Sayler, 2020, p. 5)

By implementing the Joint Artificial Intelligence Center (JAIC) as the coordinating focal point across the Department, the DoD is poised to implement the tactical lessons learned at the strategic level in the application and fielding of the new technologies that fall under the umbrella of "artificial intelligence." Because the JAIC is a new organization, the JAIC is limited in its ability to coordinate and provide oversight in the numerous AI pilots and Acquisitions across the Services and the DoD's Fourth Estate. In a comprehensive report that assessed the DoD's AI posture, the RAND Corporation found the following:

> The JAIC lacks the authorities to carry out its present role. At its core, the JAIC's overarching mission can be distilled to this: Scale AI and its impact across DoD. This mission and its present scope—as defined by the summary of the DoD AI strategy and the memo establishing the JAIC—are extensive, while the JAIC's current authorities are limited. In particular, the JAIC is expected to synchronize DoD AI activities and coordinate AI initiatives totaling more than $15 million annually. (Tarraf et al., 2019, p. 47)

Thus, while the DoD has improved its posture for AI to meet the requirements set forth in the Third Offset Strategy, it is uncertain whether with its limited authority, the JAIC is positioned to effectively coordinate the development and procurement of AI across the Department.

Acquisition Life Cycle & AI

Housel et al. (2019) noted that the DoD 5000-series Acquisition Life Cycle (see Figure 19) can be aligned to the generic technology investment life cycle. As depicted in Figure 20, while terminology differs between the DoD 5000 and generic technology life-cycle phases, the sequence of activities in these respective life cycles is congruent.

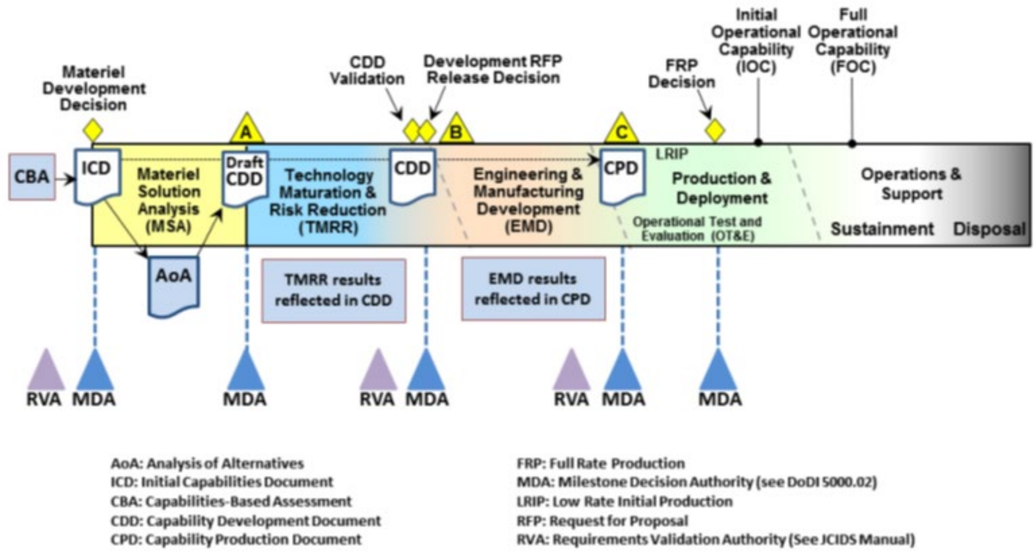


Figure 19. The 5000-Series Acquisition Life Cycle.
***Source:* DoD (2017).**

Pre-Materiel Solutions Analysis	Materiel Solutions Analysis	Technology Maturation and Risk Reduction	Engineering and Manufacturing Development	Production and Deployment	Operations and Support
-Strategic goal alignment -Pre-investment	Pre-Investment	Pre-investment	Implementation	Implementation	Post-implementation

Figure 20. Aligning the Generic and 5000-Series Life Cycles.
***Source:* Housel et al. (2019a)**

Housel et al. (2019) noted that the acquisition methodologies, including EVM, KVA, IRM, and others, may be used "concert"; however, "certain tools are more appropriate for a particular phase than others" (p. 48). Thus, the PM should "use the tools appropriately in that they provide more information for a complex environment" (Housel, et al., 2019, p. 48). Therefore, during the beginning of the life cycle such as the Pre-Material Solution Analysis, the MSA, and or the TMRR phases, methodologies such as IRM and KVA may be of more use to the PM as they provide quantitative value metrics such as a common unit assessment of the technology and risk, as opposed to EVM that only measures cost. Meanwhile, during the main implementation phases of Engineering and Manufacturing Development (EMD) and Production and Development (PD), all three metrics—EVM, KVA, and IRM—can provide useful data to assess the program. However, during the post-implementation or Operations and Support (OS) phase, KVA is more likely to provide useful data over EVM.

Acquisition Methodologies & AI Development

In a recent Congressional Research Service (CRS) on AI, it was noted that "standing DOD processes—including those related to standards of safety and performance, acquisitions, and intellectual property and data rights—present another challenge to the integration of military AI" (Sayler, 2020, p. 16). As discussed earlier in the report, the difficulty of traditional Acquisition methods in relation to emerging technology such as AI is that the speed of the development and deployment of these technologies often outpaces the Defense Acquisition System. When compared to the commercial sector, the DoD process for developing and fielding its AI systems has a mismatch when compared to the timelines for its requirements. This is likely because there is a stark contrast with "the pace of commercial innovation and DoD's acquisition process" (Ilachinski, 2017, p. 1). This is because it takes an average of 91 months, or 7 and a half years, to go from the Analysis of Alternatives (AoA) to the Initial Operational Capability (IOC; Ilachinski, 2017). The Defense Science Board (2018) found that the DoD's timeline for fielding systems is multiple times longer than the commercial sector, which uses an iterative process to field AI systems in approximately 6 months. As noted in the RAND Report that assessed the DoD's posture for AI,

> Our starting point at the onset of our study was the DoD model of technology development, procurement, fielding, and sustainment, giving rise to two dimensions of posture assessment related to technologies: advancement and adoption. However, as we carried out our study, it became clear that this model is not valid for AI, owing to the spiral nature of AI technology development. (Tarraf et al., 2019, p. 51)

Because of the challenges of procuring and fielding AI systems, the CRS Report reached the conclusion that the "DoD may need to continue to adjust its acquisitions process to account for rapidly evolving technologies such as AI" (Sayler, 2020, p. 17). Similarly, the RAND Report also noted that the DoD should utilize and adapt acquisition approaches that are "appropriate for the technology" (Tarraf et al., 2019).

One way that the DoD can adjust or adapt its acquisition process and/or approach is to adapt and utilize the Acquisition methodologies based on the complexity of the development. As noted by Ilachinski (2017) and the Defense Science Board (2018), AI systems development processes tend to be iterative in their approach. Thus, the need for rapid development and prototyping is likely to be utilized more than in traditional Acquisition Life Cycles (Tarraf et al., 2019). Despite being iterative in nature, however, the techniques recommended by Housel et al. (2019) of using EVM, KVA, IRM, or a combination of various methodologies can also be applied to the Acquisition of AI systems as shown in Table 1.

Table 1: ***Selection of Acquisition methodologies based on complexity of development***

Complexity of Development	EVM	KVA	IRM	
Mature Technologies	X	X	X	
Iterative Development		X	X	
Complex AI Systems		X	X	X
Non-Complex AI Systems	X	X	X	

Note. Iterative Development is for AI systems that are either prototypes or have rapid generational development cycles (i.e., versions).

For the mature technologies and/or non-complex AI systems, because there is typically less risk in these life cycles, the PM can follow normal EVM methodologies to monitor the progress of the development. However, the PM can also choose to utilize KVA to assess the AI systems with its common unit of value. The iterative development process was needed for rapid development cycles and prototypes such as the development process used for Project Maven, the AI-enhanced analytics used to hunt terrorists. Lieutenant General Jack Shanahan, the PM of Project Maven and first director of the JAIC noted the following about this development process:

> Shanahan characterized the initial deployment this month as "prototype warfare"—meaning that officials had tempered expectations. Over the course of about eight days, the team refined the algorithm, six times. "This is maybe one of our most impressive achievements is the idea of refinement

> to the algorithm," Shanahan said. Think of it as getting a new update to a smartphone application every day, each time improving its performance. (Weisgerber, 2017)

With a rapid spiral development process for a program like Project Maven, tools such as KVA and IRM can assess the risk of the program for decision-making, as well as provide a way to value the AI system in means beyond cost over time, as in the case of EVM.

In 2017, the DoD began the process of updating the aging Defense Travel System (DTS). For the modernization process, the DoD sought to expand the functionality of the current system as well as develop a system with "enhanced capabilities like mobile applications, business intelligence, artificial intelligence, and machine learning to automatically audit 100 percent of expense reports for fraud and compliance, and a traveler risk management and safety communication solution" (National Defense Transportation Association, 2020). With a complex system such as DTS that consists of a modernization of the current system and development of future capabilities, tools such as KVA and IRM would be useful to the PM, more so than simply performing EVM against baseline costs.

In summary, this section provided a brief overview of AI by discussing the definition, an overview of AI from its historical foundation to the current environment in the realm of international security, relating the role that AI technologies play in the Third Offset Strategy. Furthermore, this section discussed the challenge of AI based on the complexity of the systems, the risk of new technologies and inherent mistrust, as well as the DoD's current posture for AI. Finally, the section discussed how the Acquisition methodologies of EVM, KVA, and IRM can be used in conjunction with EVM to monitor AI Acquisition programs. The next section will discuss in more detail how these methodologies can be used to monitor and assess AI programs.

Methodologies

Earned Value Management

Benefits

EVM, when properly implemented, provides an early warning of potential issues within a program and a forecast of the total cost and schedule requirements (Fleming & Koppelman, 2010). Research has shown that the CPI (cost performance index) stabilizes within a 10% range after a project reaches the 20% completion point (Christensen, 1998). The 20% stabilization point holds true across various contract types, programs, and services (Fleming & Koppelman, 2010). This early indication provides project managers a reasonable prediction of the project's final costs (Fleming & Koppelman, 2010). The low range of the cost overrun is the current cost plus the remaining scheduled cost (Fleming & Koppelman, 2010). The high end of the costs is the budgeted cost divided by the CPI, which is a more accurate estimate unless extenuating circumstances caused the overrun (Fleming & Koppelman, 2010). SPI also provides an indication of future cost increases. An unfavorable SPI (schedule performance index) indicates spending will grow larger than initially planned to reduce the schedule variance (Christensen, 1998).

Each component has a cost and schedule associated with its completion that feeds into the overall completion of the project. Since the baselines are established prior to work beginning on any subcomponents, management can create well in advance a detailed timeline with the expected cost necessary to achieve the tasks. Having such a thorough plan allows the PM to focus on the areas that are reporting discrepancies in their CPI and SPI, without the benefits of a quantitative **value** metric, rather than concentrating on all areas of the project. They can collaborate when there are issues with a component, trusting the aspects of a project that are on schedule and on budget to remain within performance projected cost and schedule parameters. Instead of actively attempting to control all parts of the project, they can spend their limited time correcting cost and schedule issues via those identified areas of the project that have unfavorable SV and CV. The key metrics lead PMs to manage by exception.

Using a WBS to assign, track, and complete tasks facilitates the reporting of a project's cost and schedule performance, but not its value-to-cost ratio. The use of the existing standard EVM approach facilitates communication between the PM and the contractor with regard to cost and schedule. Since the individual components of a project are broken down into small subsections that must be completed, management can more easily communicate with contractors concerning discrepancies with subcomponents. This breakdown allows a PM to assign resources and facilitates subsequent tracking of a system's cost and schedule progress without creating additional reporting requirements. Given that reporting on the status of a program through EVM is required by law, managing with the same system may reduce administrative requirements (Christensen, 1998). EVM gives managers a single system to track cost, completion, and project schedule performance, centralizing the control system while allowing lower level managers to oversee their sections with the same limited metrics.

Quantitative simplicity is a key attribute of the EVM methodology. The process is broken down into three main variables: cost, time, and work completed. The key metrics managers use to assess a project are simple ratios derived from these cost and schedule parameters. Determining the performance of the project is as easy as determining how much money has been spent to complete the current amount of work at a given time. The methodology readily scales from the overall project to individual components without changing reporting requirements; managers simply compile the component reports into an aggregate report. However, in IT programs, it can be more difficult to determine if the work is completed satisfactorily (i.e., did it produce expected value, and in accordance with the stated requirements), as seen in the Joint Tactical Radio System (JTRS) acquisition.

Traditional cost management approaches reflect a project's funding performance and not the true cost of the program (Fleming & Koppelman, 2010). Flexible budgets also give management the option to vary the budget for work based on the amount of work completed (Christensen, 1998). EVM differs from a common budgeting approach by including the schedule dimension in its calculations. It puts the schedule variance into a cost dollar amount, quantifying delays. The amount of money spent completing a project is important, but it does not show the complete picture without schedule performance. If an initiative is on or under budget yet is delivered 7 years later than promised, it did not perform well despite the potential cost savings. By costing scheduled time, the PM can determine how much a delay will cost, providing greater flexibility within management decisions. However, this approach does not guarantee that value has been provided at an acceptable returns level.

Challenges

EVM has proven to be an effective management system in traditional manufacturing processes. However, in fiscal year 2008, the U.S. federal IT portfolio contained 346 major IT programs worth approximately $27 billion that received a rating of "unacceptable" or were on the "Management Watch List" (Kwak & Anbari, 2012). This suggests there is significant room for improvement within the current AI acquisition process. AI projects are not developed in the same manner as manufacturing-based projects that are not highly volatile. When building a warehouse, a contractor uses the architect's blueprints to determine exactly how many bolts will be needed for each beam and how long it will take to install each item. Combining each subprocess, which consists of known cost/schedule parameters, the manager can establish a baseline with a high degree of accuracy. However, when a program involves writing code, the process is more complicated due to the need for programmers to generate elegant code. Keeping with EVM principles, the desired outputs for a highly structured program are known before work begins, and the plan to accomplish those steps can be created and mapped out prior to construction of the code. However, the time required to write, test, debug, and retest the computer code can vary significantly between projects and individuals, leading to significant variability. Using the EVM system that measures the overall progress of a program based on the linear completion of subcomponents compared to a baseline established well before work began is not an accurate assessment for IT projects that do not progress linearly. AI projects are characterized by iterative development with highly variable production processes

The requirements in large AI projects are often not well defined, leading to cost and schedule overruns. For instance, JTRS used LOCs (lines of code) to create the baseline. However, LOC can vary drastically from one programmer to another or between programming languages. While there are industry standards for reducing the number of LOCs, there is still variance between individual programmers who create the AI program. In a traditional manufacturing application of EVM, a WBS containing the task "put the nut on the bolt" would have schematics and plans associated with the task so that any qualified worker could complete the work. In an AI program, turning putting the nut on the bolt via a software program could be completed in any number of ways and still meet the specifications. This simplistic example illustrates issues that will arise when using EVM to manage a more complicated IT-based project, such as creating a software-defined waveform in the JTRS project.

One of the core principles of EVM is maintaining the baseline of a program throughout the life of the project, allowing for consistent and accurate measurement of a project's cost/schedule progress. Adding additional requirements or changing the specifications of certain components can alter the trajectory of a program entirely. While any program will need to adjust when scope creep occurs, EVM is particularly ill-suited to make these modifications midstream for AI type projects. If new requirements are added to a program or existing requirements changed after the baseline is established, especially after work commences, the baseline is no longer useful as a point of comparison. More often than not, there will be additional costs and longer timelines required to fulfill the new specifications. This will change the ACWP while leaving the BCWS the same. As the actual costs increase and schedule expands compared to the baseline, the main performance metrics (CV, CPI, SV, and SPI) are negatively impacted, even if the newly added requirements add serendipitous value or miss value-adding opportunities but complete the project on time and on budget. EVM does not function well in projects with volatile scope or requirements.

When developing JTRS, much of the software had never been written or created, calling on developers to estimate the cost and schedule for the various components of a complex project with no comparable histories for comparison. To account for the possibility that a project might not be completed on time, the engineer or PM may add an additional week to something originally thought to take 3 weeks, accounting for some risk of the unknown. Should overruns occur in other areas, PMs often remove this risk factor from their calculations by reallocating scheduled time to improve their performance compared to the baseline. This omission eliminates much of the risk mitigation the program originally established, eventually leading to further discrepancies to original CV and SV estimates. However, when creating a complex product based on the assumption of future technological advancements, as in JTRS, there is a high degree of risk. Presuming all tasks will be completed on schedule without additional schedule risk mitigation is a misguided assumption to make after previous components of the project have already fallen behind schedule.

EVM functions well as a tool to monitor the cost and schedule of programs developed in waterfall or parallel design methods. In alternative design methodologies, such as agile or iterative designs, EVM does not provide the same level of flexibility to project managers. EVM does not require a particular development approach and can be used in any system that uses a baseline plan with specific schedule and cost (DoD, 2018). Techniques such as agile EVM and scrum attempt to use a more iterative approach typical of the software development process but are still using the same EVM concepts with tasks broken into various work packages. By assigning a budget and timeline to specific features within an AI initiative, PMs may use EVM to oversee progress. As each feature is completed, value is earned and the EVM metrics are updated. Within the specific features and components of a program, however, EVM is not as useful. This is because the specifics of writing computer code often are not well-structured as would be the case in a physical product project.

If two AI capabilities written in parallel work individually, it is still possible there will be issues when combining the features in the final project output. To reduce this risk, PMs feel forced to add additional time in their schedule. This can be done by scheduling multiple increments that are planned, designed, coded, tested, and demonstrated (DoD, 2018). While this is a viable way to use EVM in agile design, it is not as accurate or precise as an EVM program in a brick-and-mortar, physical product type project, which may explain, in spite of the added agility, why AI projects are notoriously over budget and over schedule.

For EVM metrics to be accurate the cost and schedule must be accurate to a great level of detail. WBSs for multibillion-dollar programs often cover all items from the strategic overview down to small tasks within the project. Creating such a detailed list of requirements for each component within an AI system is problematic. While the EVM approach may provide a simple cost and schedule estimates, the initial setup process to establish the baseline can be immense and highly inaccurate for AI projects.

Knowledge Value Added (KVA)

Benefits

KVA is an objective, quantifiable method to measure the value produced by a system and the subprocesses within the system. The value measurements of each process use ratio scale numbers, allowing analysts to compare the resulting ratios with the ratios from other subprocesses to determine their relative effectiveness. KVA converts the outputs of all processes into common value units allowing a standard productivity performance ratio across all processes. PMs can determine the value generated from the human component against the value added by IT processes. Because of the scales, PMs can use these measurements to develop useful ratios in their analysis of the program's performance. Productivity ratios such as return on knowledge (ROK) (i.e., the common units outputs of a process) are divided by the process cost required to produce the outputs, and for ROI estimates, the ratio is monetized outputs minus cost divided by cost. The ROKs and ROIs, which are always 100% correlated, give managers information about the amount of value a process generates compared to the amount of money spent to create the value. Unlike any other methodology, KVA assigns these figures to both the process and subprocesses rather than only the firm as a whole (as is done in standard generally accepted accounting practice metrics used in standard financial ratios).

Conducting an analysis of a program using KVA will give a PM a clearer picture of the value of the operational components of the program. While organizations likely have cost/schedule metrics used to determine the performance of a project or operation, ROK will give them additional value-based information to improve their management decisions. PMs can determine the relative projected baseline value of the components that comprise the program. Knowing a particular job or subprocess gives the same output value as a different process but at a different cost may provide context for understanding the variations in the performance of the overall system. This, in turn, gives experienced managers the information needed to allocate resources to specific components of a program that need improvement or should be utilized more frequently, resulting in greater value added. It also allows projections of the potential value added of a capability of an AI system that was not originally in the project plan.

While a KVA analysis can provide information that will help manage the course of a program or project, it does not require significant changes to organizational structure or reporting processes to do so. The evaluation can be conducted during normal reporting practices without introducing complicated new metrics into the system. The learning time, process instruction (e.g., WBS can be a surrogate for this method), or the binary query method are all based on information that should be available within the project description, requirements documents. A small amount of hands-on measurement may be required to verify the accuracy of the given data. As such, the analysis can be done quicker than the other common assessment methodologies (e.g., activity-based costing), giving PMs access to actionable information more rapidly.

Challenges

KVA will give analysts a quantifiable, ratio-scale number for the value of the components that produce the outputs of the subprocesses. However, it does this only with processes that consist of known a priori outputs. The intangible items, such as creativity and imagination, that occur within the human brain cannot be quantified with this method, or any other method for that matter. In fact, no current system is able to accurately quantify these types of intangibles within a process because there is no algorithm for creativity. These factors are not common to the average user and as such, cannot be defined via any of the KVA methods—learning time, binary query, or process description—because the creativity process cannot be learned or described algorithmically. Once creativity has been employed to generate an AI capability, it is possible to describe its productivity algorithmically using KVA. KVA assigns the current value of a process, but it cannot predict the value of potential future additional outputs until they are describable in one of the methods used in the KVA methodology.

Although KVA will provide ratio-scale numbers to aid in evaluating processes within a program, the ratios are often only valid for comparisons within the same project.

Benchmarking the raw numbers with other organizations or with different divisions in the same organization will provide a usable comparative performance assessment benchmark for expected ROK performance. Regardless of the language of description for outputs, the resulting ROK, ROI measures will be comparable just as they are among organizations (i.e., for profits and not for profits) that produce different products or services. Because these descriptions of output are in common units, they can be treated as a value constant across all processes where only the amount of outputs determines the value of a given component subprocess or core process. The final results of any properly conducted analysis will return the comparable ROK, ROI estimates that is the ultimate goal of KVA.

Integrated Risk Management

All organizations depend heavily on project planning tools to forecast when various projects will be complete. Completing projects within specified times and budgets and a given value is critical to facilitate smooth organizational operations. In our high-technology environment, many things can impact schedule. Technical capabilities can often fall short of expectations. Requirements may be insufficient in many cases and need further definition. Tests can bring surprising results—good or bad. A whole host of other reasons can lead to cost increases, schedule slips, and value variability. On rare occasions, we may run into good fortune, and the schedule can be accelerated without harm to the productivity of the project outcome.

Project schedules are inherently uncertain, and change is normal. Therefore, we should expect changes and find the best way to deal with them. So why do projects always take longer than anticipated? One reason is inaccurate schedule estimating. The following discussion presents a description of shortcomings in the traditional methods of schedule estimation and how simulation and advanced analytics can be applied to address these shortcomings.

It is important to first understand the Integrated Risk Management (IRM) process and how the techniques involved are related in a risk analysis and risk management context. This framework comprises eight distinct phases of a successful and comprehensive risk analysis implementation, going from a qualitative management screening process to creating clear and concise reports for management. The process was developed by the author (Dr. Mun) based on previous successful implementations of risk analysis, forecasting, real options, valuation, and optimization projects both in the consulting arena and in industry-specific problems. These phases can be performed either in isolation or together in sequence for a more robust integrated analysis.

We can segregate the process into the following eight simple steps (Mun, 2016):

1. Qualitative Management Screening
2. Forecast Predictive Modeling
3. Base Case Static Model
4. Monte Carlo Risk Simulation
5. Real Options Problem Framing
6. Real Options Valuation and Modeling
7. Portfolio and Resource Optimization
8. Reporting, Presentation, and Update Analysis

Qualitative Management Screening

Qualitative management screening is the first step in any IRM process. Management has to decide which projects, assets, initiatives, or strategies are viable for further analysis, in accordance with the firm's mission, vision, goal, or overall business strategy, all of which may include market penetration strategies, competitive advantage, technical, acquisition, growth, synergistic, or globalization issues. That is, the initial list of projects should be qualified in terms of meeting management's agenda. Often the most valuable insight is created as management frames the complete problem to be resolved. This step is where the various risks to the firm are identified and flushed out.

Forecast Predictive Modeling

The future is then forecasted using time-series analysis or multivariate regression analysis if historical or comparable data exist. Otherwise, other qualitative forecasting methods may be used (subjective guesses, growth rate assumptions, expert opinions, Delphi method, etc.). In a financial context, this is the step where future revenues, sale price, quantity sold, volume, production, and other key revenue and cost drivers are forecasted. Examples of methods include time-series, nonlinear extrapolation, stochastic process, ARIMA, multivariate regression forecasts, fuzzy logic, neural networks, econometric models, GARCH, and so on.

Base Case Static Model

Whether for individual or multiple projects under evaluation, for each project that passes the initial qualitative screens, a discounted cash-flow model is created (KVA analysis, using the market comparables approach can be used to monetize value for this model). This model serves as the base case analysis where a net present value is calculated for each project, using the forecasted values in the previous step. This net present value is calculated using the traditional approach of modeling and forecasting revenues and costs and discounting the net of these revenues and costs at an appropriate risk-adjusted rate. The return on investment and other profitability, cost-benefit, and productivity metrics are generated here.

Monte Carlo Risk Simulation

Because the static discounted cash flow produces only a single-point estimate result, there is often little confidence in its accuracy given that future events that affect forecast cash flows are highly uncertain. To better estimate the actual value of a particular project, Monte Carlo risk simulation should be employed next. Usually, a sensitivity analysis is first performed on the discounted cash-flow model; that is, setting the net present value as the resulting variable, we can change each of its precedent variables and note the change in the resulting variable. Precedent variables include revenues, costs, tax rates, discount rates, capital expenditures, depreciation, and so forth, which ultimately flow through the model to affect the net present value figure. By tracing back all these precedent variables, we can change each one by a preset amount and see the effect on the resulting net present value. A graphical representation can then be created, which is oftentimes called a tornado chart because of its shape, where the most sensitive precedent variables are listed first, in descending order of magnitude. Armed with this information, the analyst can then decide which key variables are highly uncertain in the future and which are deterministic. The uncertain key variables that drive the net present value and, hence, the decision are called *critical success drivers*. These critical success drivers are prime candidates for Monte Carlo simulation. Because some of these critical success drivers may be correlated—for example, operating costs may increase in proportion to quantity sold of a particular product, or prices may be inversely correlated to quantity sold—a correlated Monte Carlo simulation may be required. Typically, these correlations can be obtained through historical data. Running correlated simulations provides a much closer approximation to the variables' real-life behaviors.

Real Options Problem Framing

After quantifying risks in the previous step, the question now is, what is next? The risk information obtained somehow needs to be converted into *actionable intelligence*. Just because risk has been quantified to be such and such using Monte Carlo simulation, so what, and what do we do about it? The answer is to use real options analysis to hedge these risks, to value these risks, and to position yourself to take advantage of the risks. The first step in real options is to generate a strategic map through the process of framing the problem. Based on the overall problem identification occurring during the initial qualitative management screening process, certain strategic optionalities would have become apparent for each particular project. The strategic optionalities may include among other things, the option to expand, contract, abandon, switch, choose, and so forth. Based on the identification of strategic optionalities that exist for each project or at each stage of the project, the analyst can then choose from a list of options to analyze in more detail. Real options are added to the projects to hedge downside risks and to take advantage of upside swings.

Real Options Valuation and Modeling

Through the use of Monte Carlo risk simulation, the resulting stochastic discounted cash-flow model will have a distribution of values. Thus, simulation models, analyzes, and quantifies the various risks and uncertainties of each project. The result is a distribution of the NPVs and the project's volatility. In real options, we assume that the underlying variable is the future profitability of the project, which is the future cash-flow series. An implied volatility of the future free cash flow or underlying variable can be calculated through the results of a Monte Carlo simulation previously performed. 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). In addition, the present value of future cash flows for the base case discounted cash-flow model is used as the initial underlying asset value in real options modeling. Using these inputs, real options analysis is performed to obtain the projects' strategic option values.

Portfolio and Resource Optimization

Portfolio optimization is an optional step in the analysis. If the analysis is done on multiple projects, management should view the results as a portfolio of rolled-up projects because the projects are in most cases correlated with one another, and viewing them individually will not present the true picture. As firms do not only have single projects, portfolio optimization becomes crucial. Given that certain projects are related to others, there are opportunities for hedging and diversifying risks through a portfolio. Because firms have limited budgets, as well as time and resource constraints, while simultaneously having requirements for certain overall levels of returns, risk tolerances, and so forth, portfolio optimization takes into account all these to create an optimal portfolio mix. The analysis will provide the optimal allocation of investments across multiple projects.

Reporting, Presentation, and Update Analysis

The analysis is not complete until reports can be generated. Not only must the results be presented, but the process should also be shown. Clear, concise, and precise explanations transform a difficult black box set of analytics into transparent steps. Management will never accept results coming from black boxes if they do not understand where the assumptions or data originate and what types of mathematical or financial massaging takes place. Risk analysis assumes that the future is uncertain and that management has the right to make midcourse corrections when these uncertainties become resolved or risks become known; the analysis is usually done ahead of time and thus, ahead of such uncertainty and risks. Therefore, when these risks become known, the analysis should be revisited to incorporate the decisions made or revising any input assumptions. Sometimes, for long-horizon projects, several iterations of the real options analysis should be performed, where future iterations are updated with the latest data and assumptions.

Understanding the steps required to undertake the IRM process is important because it provides insight not only into the methodology itself but also into how it evolves from traditional analyses, showing where the traditional approach ends and where the new analytics start.

Applying IRM: Schedule and Risk Management

Traditional schedule management typically starts with a list of tasks. Next these tasks are put in order and linked from the predecessor to successor for each task. They are typically displayed in either a Gantt chart form or a network. For our discussion in this section, we concentrate on the network diagram. The duration for each task within the network is then developed. The estimated duration for each task is given a single-point estimate, even though we know from experience that this estimate should be a range of values. Therefore, the first mistake is using a single-point estimate. In addition, many people who provide duration estimates try to put their best foot forward and give an optimistic or best-case estimate. If we assume that the probability of achieving this best-case estimate for one task is 20%, then the likelihood of achieving the best case for two tasks is merely 4% (20% of 20%), and three tasks yields only 0.8%. Within a real project with many more tasks, there is only an infinitesimal chance of making the best-case schedule. Once the task duration estimates have been developed, the network is constructed and the various paths through the network are traced. The task durations are summed along each of these paths, and the one that takes the longest is identified as the critical path.

Figure 21 illustrates a network and critical path using sample data (Mun, 2016a). The sum of task durations along the critical path is listed as the project completion date. In Figure 18, there are four paths through the network from beginning to end. The shortest/quickest path is Tasks 1-2-3-10-11 with a total duration of 22 days. The next shortest path is tasks 1-7-8-9-10-11 at 34 days, and then path 1-4-5-6-10-11 at 36 days.

Finally, the path 1-4-8-9-10-11 takes the longest at 37 days and is the critical path for this network.

So, let us assume that this network of tasks is our part of a larger effort and some other effort upstream of ours has overrun by a day. Our boss has asked us to shorten our schedule by 1 or 2 days to get the overall effort back on track. Traditional schedule management has one target: shorten the longest duration item in the critical path. Another approach is to shorten every task on the entire critical path. Because the first technique is more focused, more prone to success, and creates fewer conflicts on our team, let us assume that we will use that one. Hence, we will want to reduce Task 8 from 10 days to 9 days to shorten our schedule and we will satisfy our boss or our customer. Let us leave the traditional methodology at this stage feeling satisfied with our efforts, but curious about exploring alternatives. The next step is to explore simulation and risk analytics to enhance the management of the project. Specifically, we will be employing Monte Carlo risk simulations on each of the tasks' projected budget and schedule, resulting in a probabilistic and risk profile view of the entire network's cost and schedule.

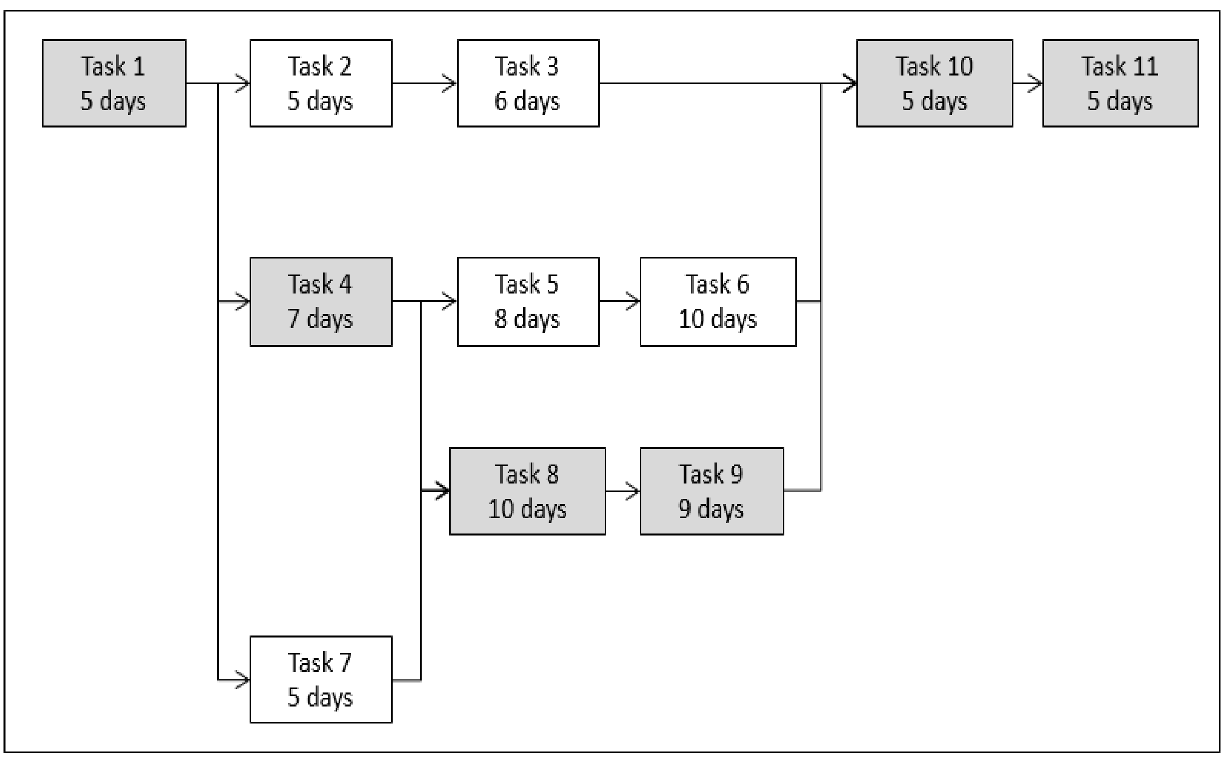


Figure 21. Complex network task

Probabilistic Schedule Management

If we agree that task durations can vary, then that uncertainty should be considered in schedule models. A schedule model can be developed by creating a probability distribution for each task, representing the likelihood of completing the particular task at a specific duration. Monte Carlo simulation techniques can then be applied to forecast the entire range of possible project durations.

A simple triangular distribution is a reasonable probability distribution to use to describe the uncertainty for a task's duration. It is a natural fit because if we ask someone to give a range of duration values for a specific task, they usually supply two of the distribution's elements: the minimum duration and the maximum duration. We need only ask or determine the most likely duration to complete the triangular distribution. The parameters are simple, intuitively easy to understand, and readily accepted by customers and bosses alike. Other more complex distributions could be used such as the Beta or Weibull but little, if anything, is gained because the determination of the estimated parameters for these distributions is prone to error and the method of determination is not easily explainable to the customer or boss.

To get the best estimates, we should use multiple sources to get the estimates of the minimum, most likely, and maximum values for the task durations. We can talk to the contractor, the project manager, and the people doing the hands-on work and then compile a list of duration estimates. Historical data can also be used, but with caution because many efforts may be similar to past projects but usually contain several unique elements or combinations. We can use Figure 22 as a guide. Minimum values should reflect optimal utilization of resources. Maximum values should consider substantial problems, but it is not necessary to account for the absolute worst case where everything goes wrong, and the problems compound each other. Note that the most likely value will be the value experienced most often, but it is typically less than the median or mean in most cases. For our example problem, shown in Figure 21, the minimum, most likely, and maximum values given in Figure 23 will be used. We can use Risk Simulator software to set input assumptions to create triangular distributions based on these minimum, most likely, and maximum parameters. The column of dynamic duration values shown in Figure

23 was created by taking one random sample from each of the associated triangular distributions.

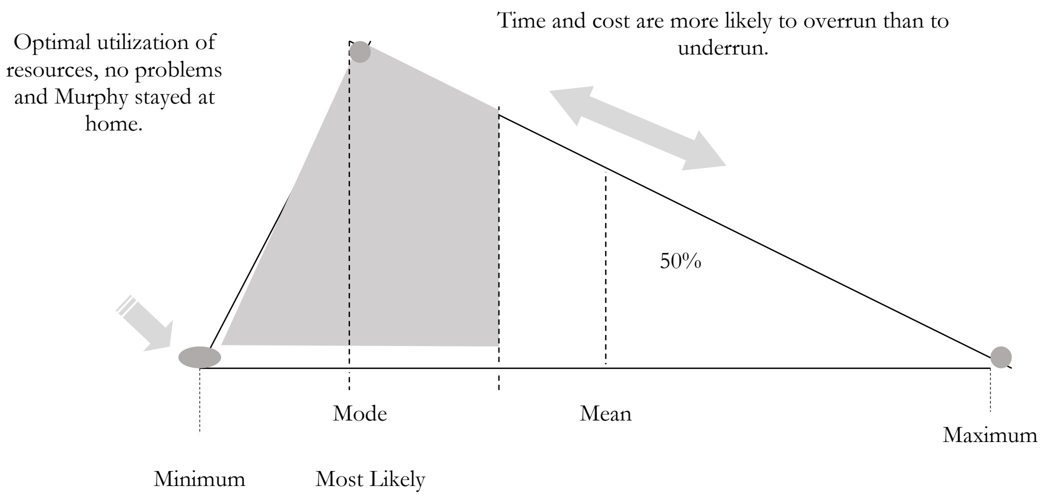


Figure 22. Triangular distribution

After the triangular distributions are created, the next step is to use the schedule network to determine the paths. For the example problem shown Figure 21, there are four paths through the network from beginning to end. These paths are shown in Figure 24 with their associated durations. The overall schedule total duration is the maximum of the four paths. In Risk Simulator, we would designate that value as an Output Forecast. In probabilistic schedule analysis, we are not concerned with the critical path/near-critical path situations because the analysis automatically accounts for all path durations through the calculations.

We can now use Risk Simulator to run a Monte Carlo simulation to produce a forecast for schedule duration. Figure 25 shows the results for the example problem. Let us return to the numbers given by the traditional method. The original estimate stated the project would be complete in 37 days. If we use the left-tail function on the forecast chart, we can determine the likelihood of completing the task in 37 days based on the Monte Carlo simulation. In this case, there is a mere 8.27% chance of completion within the 37 days. This result illustrates the second shortcoming in the traditional method: Not only is the point estimate incorrect, but it puts us in a high-risk overrun situation before the work has even started! As shown in Figure 24, the median value is 38.5 days. Some industry standards recommend using the 80% certainty value for most cases, which equates to 39.5 days in the example problem.

Task #	Task Name	Min	Likely	Max	Point Estimate
1	Stakeholder Analysis	4.5	5	6	5
2	Objectives Hierarchy	4.5	5	6	5
3	Decision Metrics Development	5.5	6	7	6
4	Functional Analysis	6	7	9	7
5	Primary Module Requirements	7	8	10	8
6	Primary Module Development	9	10	13	10
7	Secondary Module Functional Analysis	4.5	5	6	5
8	Secondary Module Requirements	9	10	12	10
9	Secondary Module Development	8	9	10	9
10	Trade Studies	2.5	3	4	3
11	Final Development Specification	2.5	3	4	3

Figure 23. Range of task durations

Path 1	Time 1	Path 2	Time 2	Path 3	Time 3	Path 4	Time 4
1	5.78	1	5.78	1	5.78	1	5.78
2	4.79	4	7.78	4	7.78	7	5.20
3	6.16	5	9.22	8	10.05	8	10.05
10	3.33	6	10.12	9	9.40	9	9.40
11	3.76	10	3.33	10	3.33	10	3.33
		11	3.76	11	3.76	11	3.76
Totals	23.82		39.99		40.10		37.52
Overall Total Schedule			40.10	(Max of all the totals)			

Figure 24. Paths and durations

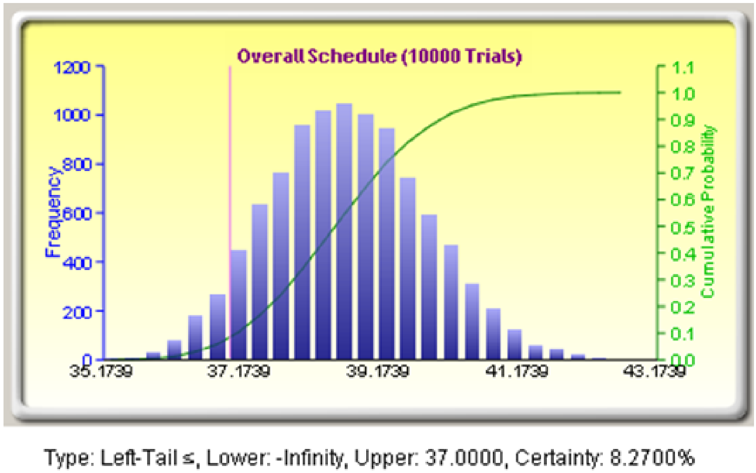


Figure 25. Simulation results

Now let us revisit the boss's request to reduce the whole schedule by 1 day. Where do we put the effort to reduce the overall duration? If we are using probabilistic schedule management, we do not use the critical path; so, where do we start? Using Risk Simulator's Tornado Analysis and Sensitivity Analysis tools, we can identify the most effective targets for reduction efforts. The tornado chart (Figure 26) identifies the most influential variables (tasks) to the overall schedule. This chart identifies the best targets to reduce the mean/median values.

We cannot address the mean/median without addressing the variation, however. The Sensitivity Analysis tool shows what variables (tasks) contribute the most to the variation in the overall schedule output (see Figure 27). In this case, we can see that the variation in Task 4 is the major contributor to the variation in the overall schedule. Another interesting observation is that the variation in Task 6, a task not on the critical path, is also contributing nearly 9% of the overall variation.

In this example, reducing the schedule duration for Task 4, Task 8, and Task 9 would pay the most dividends as far as reducing the overall schedule length. Determining the underlying reasons for the substantial variation in Tasks 4, 6, and 8 would likely give better insight into these processes. For example, the variation in Task 4 may be caused by the lack of available personnel. Management actions could be taken to dedicate personnel to the effort and reduce the variation substantially, which would reduce the overall variation and enhance the predictability of the schedule. Digging into the reasons for variation will lead to targets where management actions will be most effective, much more so than by simply telling the troops to reduce their task completion time.

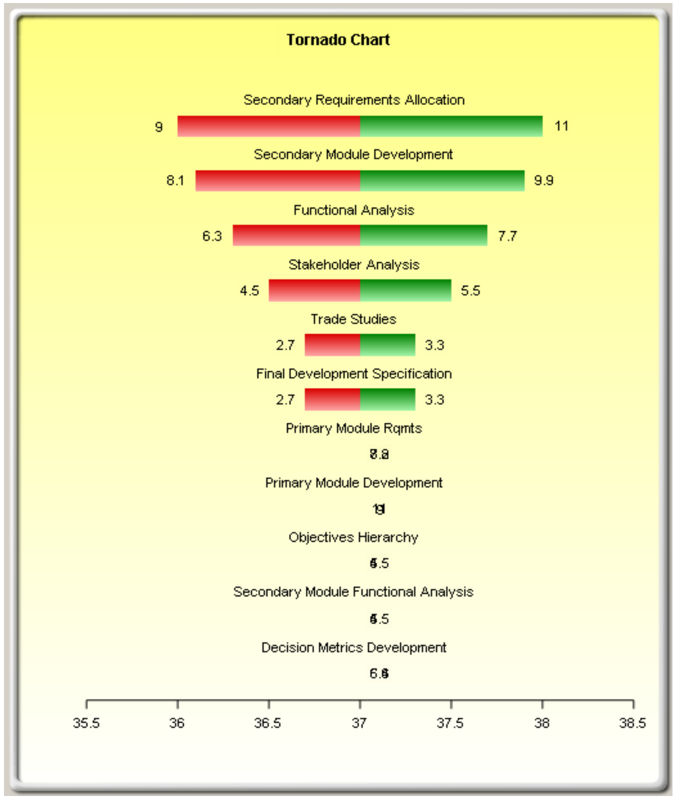


Figure 26. Tornado analysis

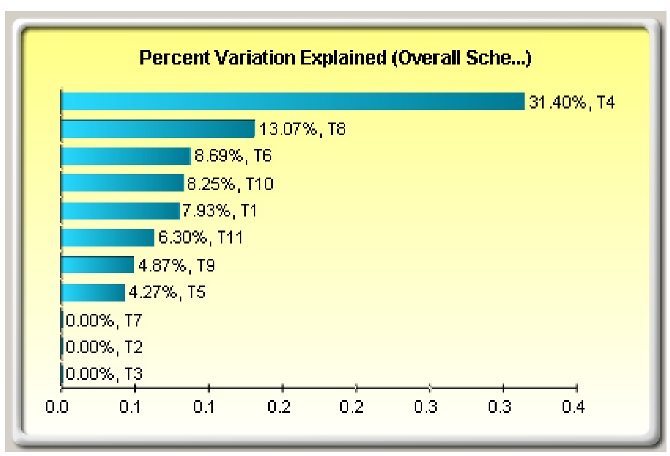


Figure 27. Sensitivity analysis

Using the network schedule model, we can also experiment to see how different reduction strategies may pay off. For example, taking 1 day out of Tasks 4, 8, and 9 under the traditional method would lead us to believe that a 3-day reduction has taken place, but if we reduce the Most Likely value for Tasks 4, 8, and 9 by 1 day and run the Monte Carlo risk simulation, we find that the median value is still 37.91, or only a 0.7-day reduction. This small reduction proves that the variation must be addressed. If we reduce the variation by 50%, keeping the original minimum and the most likely values, but reducing the maximum value for each distribution, then we reduce the median from 38.5 to 37.91—about the same as reducing the most likely values. Taking both actions (reducing the most likely and maximum values) reduces the median to 36.83, giving us a 55% chance of completing within 37 days. This analysis proves that reducing the most likely value and the overall variation is the most effective action.

To get to 36 days, we need to continue to work down the list of tasks shown in the sensitivity and tornado charts (Figures 26 and 27) addressing each task. If we give Task 1 the same treatment, reducing its most likely and maximum values, then completion within 36 days can be accomplished with a 51% certainty, and a 79.25% certainty of completing within 37 days. The maximum value for the overall schedule is reduced from more than 42 days to less than 40 days. Substantial management efforts would be needed, however, to reach 36 days at the 80% certainty level.

When managing the production schedule, use the best-case numbers. If we use the most likely values or worse yet, the maximum values, production personnel will not strive to hit the best-case numbers, thus implementing a self-fulfilling prophecy of delayed completion. When budgeting, we should create the budget for the median outcome but recognize that there is uncertainty in the real world as well as risk. When relating the schedule to the customer, provide the values that equate to the 75% to 80% certainty level. In most cases, customers prefer predictability (on-time completion) over potentially speedy completion that includes significant risk. Finally, acknowledge that the "worst case" can conceivably occur and create contingency plans to protect your organization in case it does occur. If the "worst case"/maximum value is unacceptable, then make the appropriate changes in the process to reduce the maximum value of the outcome to an acceptable level.

With traditional schedule management, there is only one answer for the scheduled completion date. Each task gets one duration estimate, and that estimate is accurate only if everything goes according to plan, which is not a likely occurrence. With probabilistic schedule management, thousands of trials are run exploring the range of possible outcomes for schedule duration. Each task in the network receives a time estimate distribution, accurately reflecting each task's uncertainty. Correlations can be entered to more accurately model real-world behavior. Critical paths and near-critical paths are automatically considered, and the output forecast distribution will accurately reflect the entire range of possible outcomes. Using tornado and sensitivity analyses, we can maximize the effectiveness of our management actions to control schedule variations and if necessary, reduce the overall schedule at high certainty levels.

Projects with Complex Tasks

In complex projects where there are nonlinear bifurcating and recombining paths (Figure 28), the cost and schedule risk modeling is more difficult to model and compute. For instance, in the *Project A* tab of the default example, we can see that after Task 1, future tasks can be run in parallel (Tasks 2, 3, and 4). Then, Tasks 3 and 4 recombine into Task 8. Such complex path models can be created by the user simply by adding tasks and combining them in the visual map as shown, using the relevant icon tools (Figure 28). The software will automatically create the analytical financial model when *Create Model* is clicked. That is, you will be taken to the *Schedule & Cost* tab and the same setup as shown previously is now available for data entry for this complex model (Figure 26). The complex mathematical connections will automatically be created behind the scenes to run the calculations so that the user will need to perform only the very simple tasks of drawing the complex network path connections.

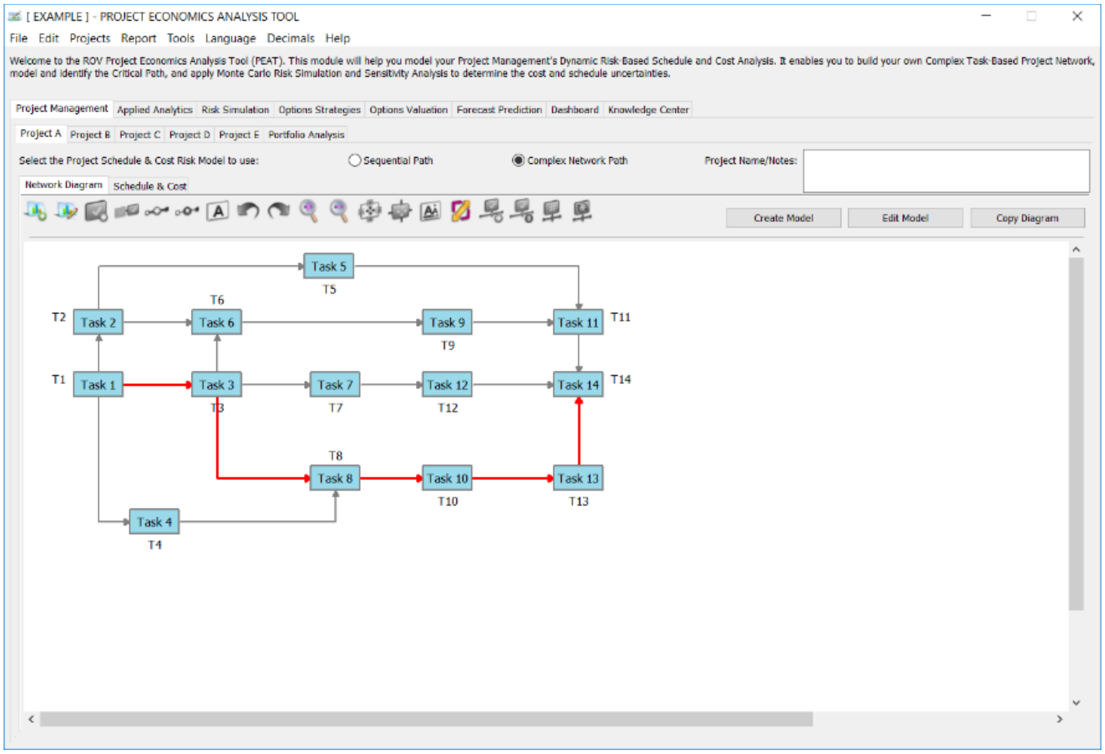
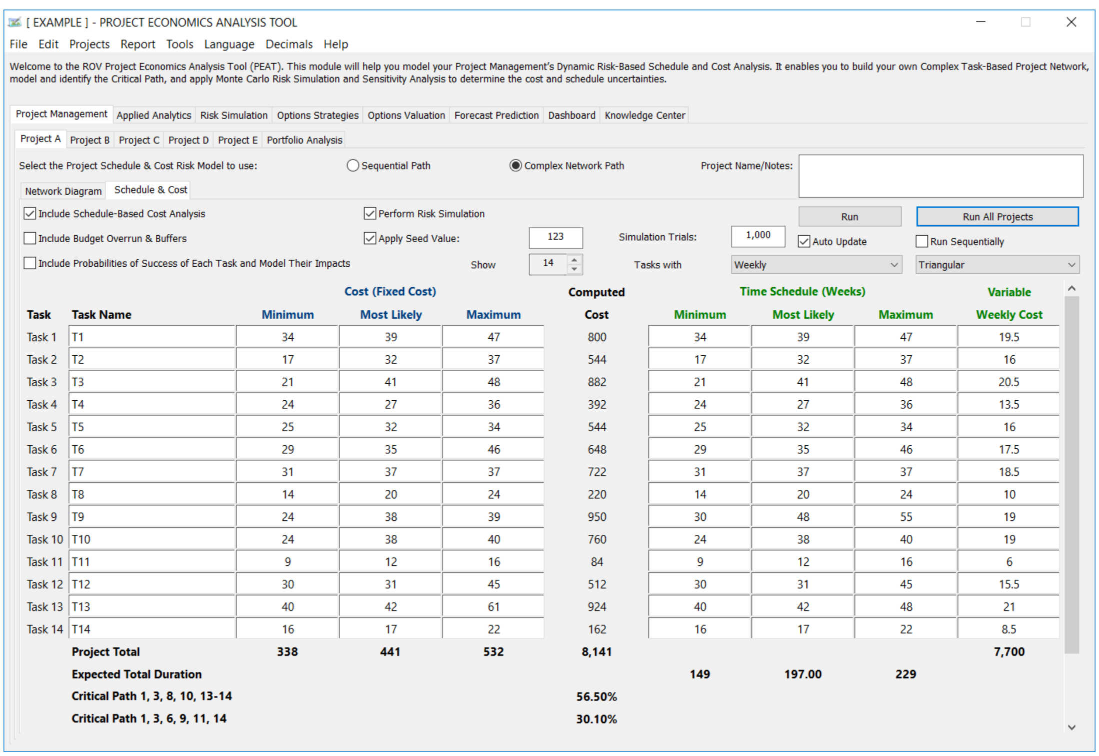


Figure 28. Complex path project management



		Cost (Fixed Cost)			Computed	Time Schedule (Weeks)			Variable
Task	Task Name	Minimum	Most Likely	Maximum	Cost	Minimum	Most Likely	Maximum	Weekly Cost
Task 1	T1	34	39	47	800	34	39	47	19.5
Task 2	T2	17	32	37	544	17	32	37	16
Task 3	T3	21	41	48	882	21	41	48	20.5
Task 4	T4	24	27	36	392	24	27	36	13.5
Task 5	T5	25	32	34	544	25	32	34	16
Task 6	T6	29	35	46	648	29	35	46	17.5
Task 7	T7	31	37	37	722	31	37	37	18.5
Task 8	T8	14	20	24	220	14	20	24	10
Task 9	T9	24	38	39	950	30	48	55	19
Task 10	T10	24	38	40	760	24	38	40	19
Task 11	T11	9	12	16	84	9	12	16	6
Task 12	T12	30	31	45	512	30	31	45	15.5
Task 13	T13	40	42	61	924	40	42	48	21
Task 14	T14	16	17	22	162	16	17	22	8.5
	Project Total	338	441	532	8,141				7,700
	Expected Total Duration					149	197.00	229	
	Critical Path 1, 3, 8, 10, 13-14				56.50%				
	Critical Path 1, 3, 6, 9, 11, 14				30.10%				

Figure 29. Complex project simulated cost and duration model with critical path

Critical Path Models (CPM) in Projects with Complex Tasks

After running the model, the complex path map shows the highlighted critical path (Figure 30) of the project, that is, the path that has the highest potential for bottlenecks and delays in completing the project on time. The exact path specifications and probabilities of being on the critical path is seen in Figure 27 (e.g., there is a 56.30% probability that the critical path will be along Tasks 1, 3, 8, 10, 13, 14).

If there are multiple projects or potential project path implementations, the portfolio view (Figure 28) compares all projects and implementation paths for the user to make a better and more informed risk-based decision. The simulated distributions can also be overlaid (Figure 29) for comparison.

Figure 31 allows users to see all projects that were modeled at a glance. Each project modeled can actually be different projects or the same project modeled under different assumptions and implementation options (i.e., different ways of executing the project), to see which project or implementation option path makes more sense in terms of cost and schedule risks. The *Analysis of Alternatives* radio button selected allows users to see each project as stand-alone (as compared to *Incremental Analysis* where one of the projects is selected as the base case and all other projects' results show their differences from the base case) in terms of cost and schedule: single-point estimate values, simulated averages, the probabilities that each of the projects will have a cost or schedule overrun, and the 90th percentile value of cost and schedule. Of course, more detailed analysis can be obtained from the *Risk Simulation | Simulation Results* tab, where users can view all of the simulation statistics and select any confidence and percentile values to show. This *Portfolio Analysis* tab also charts the simulated cost and schedule values using bubble and bar charts for a visual representation of the key results.

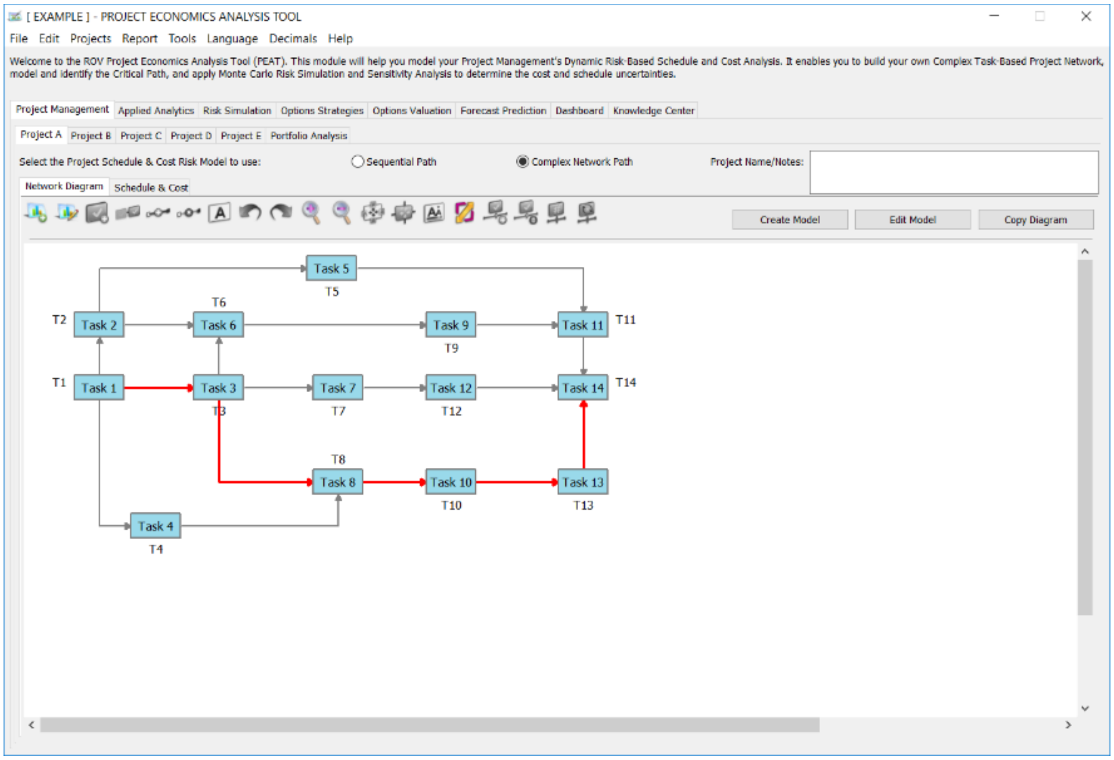


Figure 30. Complex project critical path

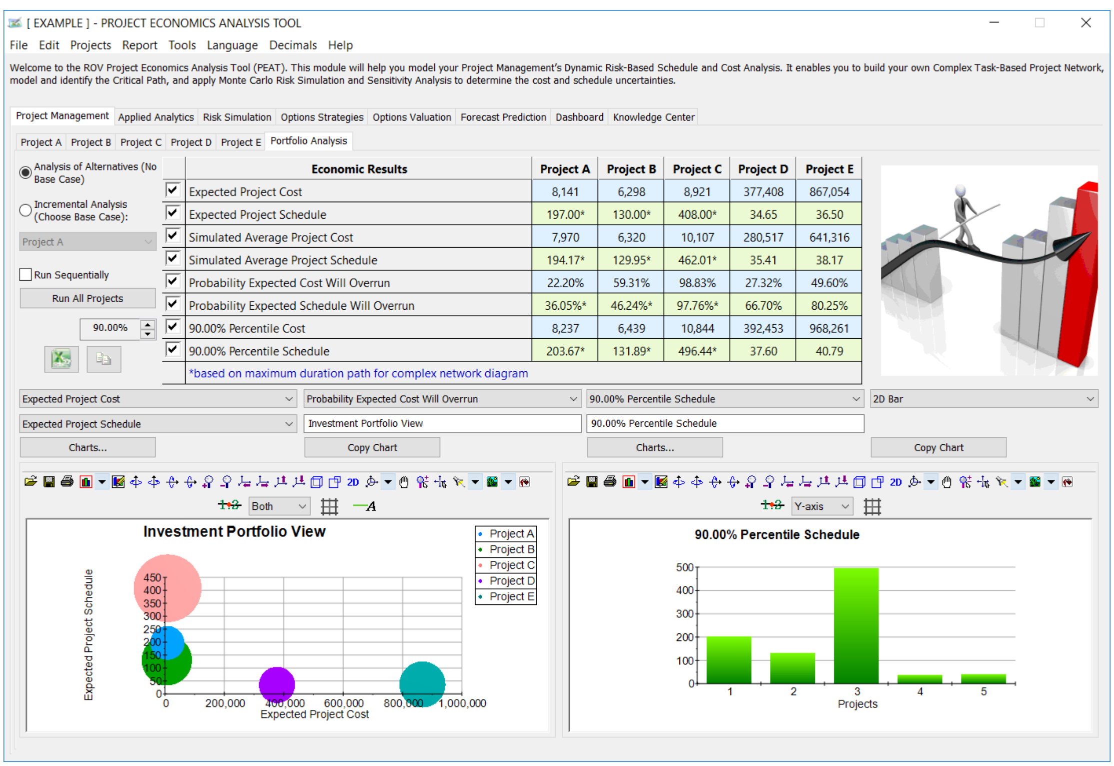


Figure 31. Portfolio view of multiple projects

Comparing and Overlaying Simulated Results

The *Overlay* chart in Figure 32 shows multiple projects' simulated costs or schedules overlaid on top of one another to see their relative spreads, location, and skew of the results. We clearly see that the project whose distribution lies to the right has a much higher cost to complete than the left, with the project on the right also having a slightly higher level of uncertainty in terms of cost spreads. Finally, Figure 33 shows an *Analysis of Alternatives* comparison of the simulated results of the projects. While Figure 29 shows the expected value of the project costs and schedule (not simulated, static, single-point estimates), Figure 33 shows the simulated results.

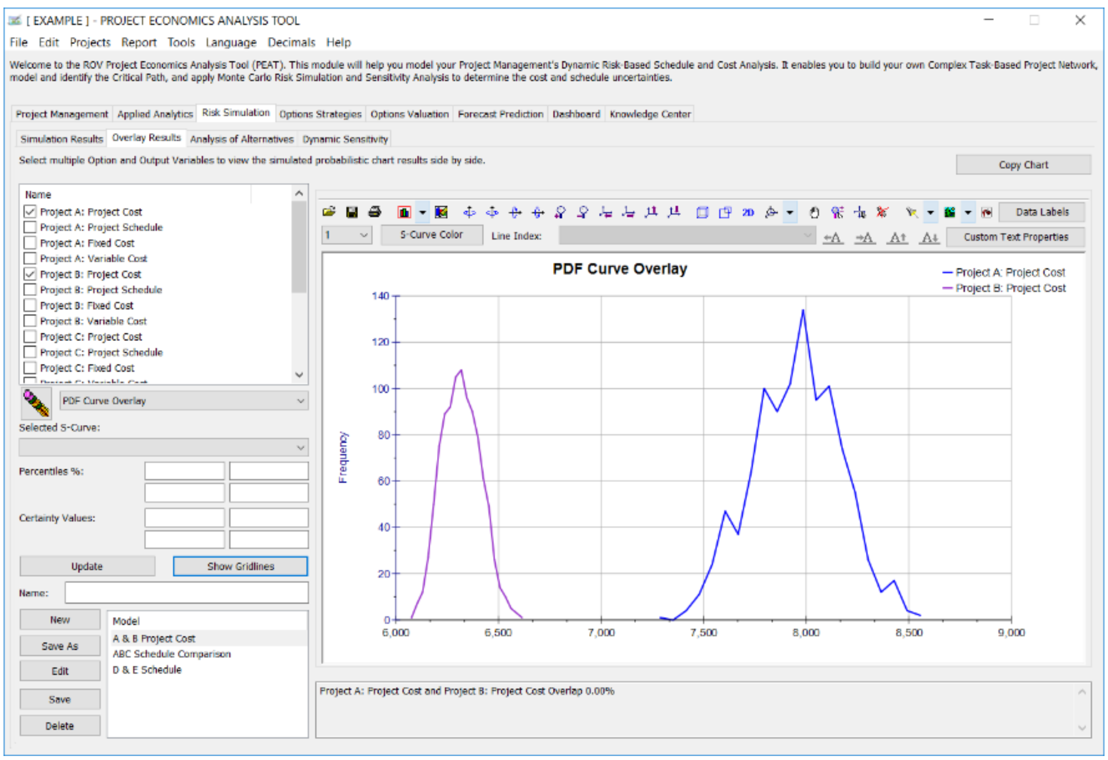
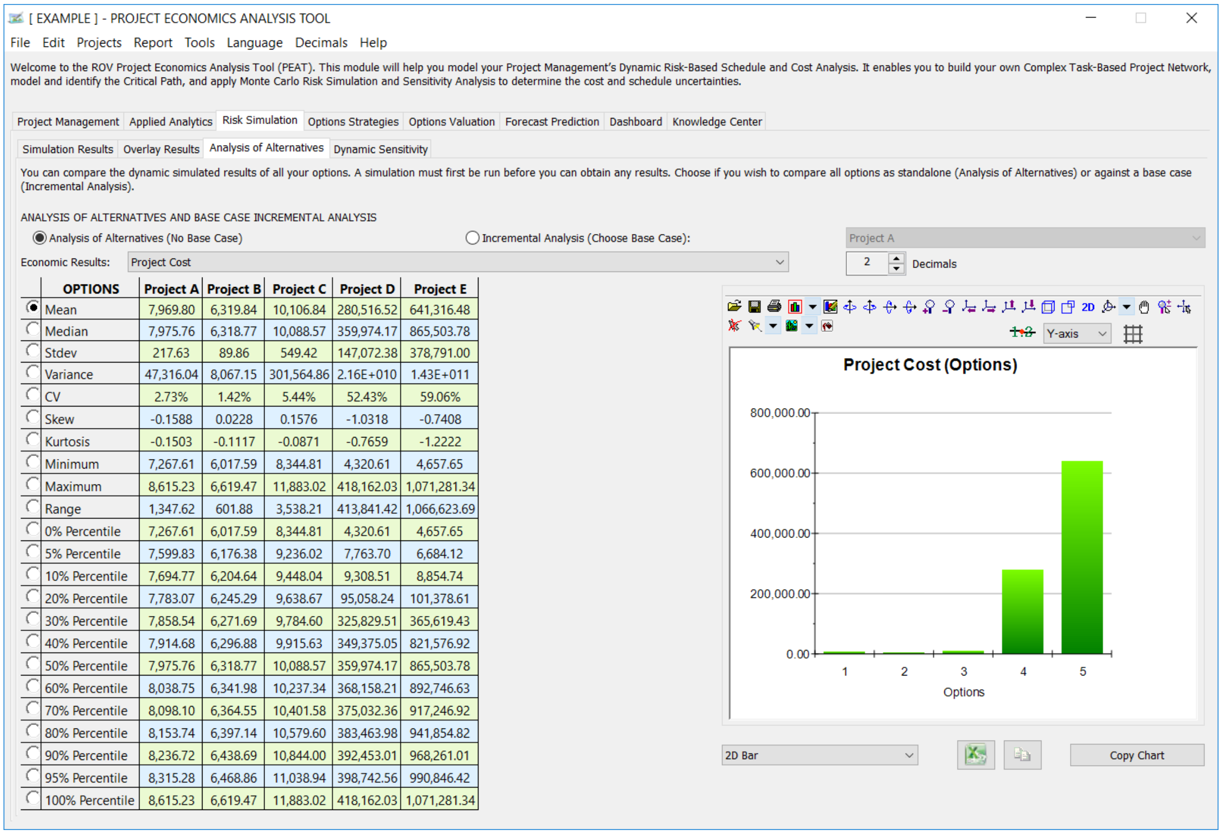


Figure 32. Overlay charts of multiple projects' cost or schedule



OPTIONS	Project A	Project B	Project C	Project D	Project E
Mean	7,969.80	6,319.84	10,106.84	280,516.52	641,316.48
Median	7,975.76	6,318.77	10,088.57	359,974.17	865,503.78
Stdev	217.63	89.86	549.42	147,072.38	378,791.00
Variance	47,316.04	8,067.15	301,564.86	2.16E+010	1.43E+011
CV	2.73%	1.42%	5.44%	52.43%	59.06%
Skew	-0.1588	0.0228	0.1576	-1.0318	-0.7408
Kurtosis	-0.1503	-0.1117	-0.0871	-0.7659	-1.2222
Minimum	7,267.61	6,017.59	8,344.81	4,320.61	4,657.65
Maximum	8,615.23	6,619.47	11,883.02	418,162.03	1,071,281.34
Range	1,347.62	601.88	3,538.21	413,841.42	1,066,623.69
0% Percentile	7,267.61	6,017.59	8,344.81	4,320.61	4,657.65
5% Percentile	7,599.83	6,176.38	9,236.02	7,763.70	6,684.12
10% Percentile	7,694.77	6,204.64	9,448.04	9,308.51	8,854.74
20% Percentile	7,783.07	6,245.29	9,638.67	95,058.24	101,378.61
30% Percentile	7,858.54	6,271.69	9,784.60	325,829.51	365,619.43
40% Percentile	7,914.68	6,296.88	9,915.63	349,375.05	821,576.92
50% Percentile	7,975.76	6,318.77	10,088.57	359,974.17	865,503.78
60% Percentile	8,038.75	6,341.98	10,237.34	368,158.21	892,746.63
70% Percentile	8,098.10	6,364.55	10,401.58	375,032.36	917,246.92
80% Percentile	8,153.74	6,397.14	10,579.60	383,463.98	941,854.82
90% Percentile	8,236.72	6,438.69	10,844.00	392,453.01	968,261.01
95% Percentile	8,315.28	6,468.86	11,038.94	398,742.56	990,846.42
100% Percentile	8,615.23	6,619.47	11,883.02	418,162.03	1,071,281.34

Figure 33. Analysis of alternatives

Benefits

The combination of several proven techniques makes IRM a valuable tool to improve the quality of information available when making decisions. Introducing dynamic Monte Carlo simulation to analysis of potential initiatives and investments illustrates the risks associated with the projects in a more realistic manner than traditional approaches. Static forecasting based on assumptions and historical performance offers a limited view of the range of a project's possible outcomes. Running thousands of simulations or more while adjusting the variables within realistic possibilities allows decision-makers to obtain a more complete picture of the uncertainty inherent within the project. Increasing the amount of relevant and accurate information managers can access will improve the quality of decisions made by the leadership team.

IRM provides a systematic approach to addressing AI investments. Following the eight steps is a straightforward process that facilitates a quantitative decision-making process. While the functions within each step are sometimes complicated and require additional training to complete, the process as a whole is clear-cut and easy to follow. Since the IRM method is well defined, it can be implemented into established procedures without a complete reengineering of established processes. Data used in traditional methods will be used in IRM and expanded to improve the scope of a project's evaluation. The real options were quantified and resulted in an output that differed from expectations. The systemic nature of IRM allows the process to be completed by different members or different teams without re-collecting data and starting from the beginning. After analysts have completed training in IRM, they should be able to continue the process from any point within the method.

Armed with the probability of certain outputs from the project, real options analysis allows managers to determine the best method to proceed with a project. Real options were presented not only at the beginning of the program with three different directions for the program to head, but also after each different phase of the chosen strategy. By writing a contract that lets an organization adjust their chosen course of action as more information becomes available, the company can minimize losses for failing programs and capitalize on initiatives that are performing well or show promise. Fortunately, within DoD acquisitions, many real options are already commonplace. The government regularly cancels contracts due to a change in fiscal policy, failure to meet specifications, or other reasons. Adding other real options to contracts would not be a completely foreign concept.

A key strength of the IRM methodology is the use of common units to make strategic decisions relate to the value of a system. By implementing KVA measurements into the static and dynamic IRM models, leadership can see a statistical range depicting the potential value of a project. The market comparable prices generated from the value analysis were used to determine the present values of the real option strategies. Most other methods use only the cost of the program to determine its effectiveness, assuming the value is inherent due to the requirements that were generated. IRM can inform decision-makers about both the costs associated with a potential investment into an initiative and the value of the said initiative in units that can be directly compared.

Challenges

While IRM is an extremely useful analytical tool, there are drawbacks to the process as well. The various techniques within the method can be difficult to master (Housel et al., 2019). It is a complicated process that requires a detailed understanding of both finance and statistics to complete a thorough analysis. While there are software tools to assist in conducting the analysis, the inputs are more complicated than simply inserting a few numbers into a program and reading an output. However, with an understanding of the basic concepts, sufficient training, and the correct tools, an analyst can generate the necessary information to allow decision-makers access to the appropriate comparative material to make an informed decision (Housel et al., 2019). The information congregated during the statistical analysis can appear daunting. For those without a strong statistics background, the simulations and their outputs seemingly come from a quantitative black box (Mun, 2016). If decision-makers do not understand why an analyst makes a recommendation, it can be easy to disregard the suggestions and rely on familiar techniques. Creating comprehensive and thorough reports for management review combined with informed presentations to alleviate concerns with the unaccustomed procedures will combat this potential issue. To take advantage of real options, these options must be considered in advance of the decision to enact any of the options. Leadership must recognize the future option when writing contracts to ensure certain options remain available. Some options, such as the option to expand, can be enacted relatively simply by developing another project based on the success of the initial investment. However, the option to abandon may not be as readily available to project managers if the contract did not include appropriate clauses. Vendors must be willing to accept potential cancellation of subcontracts when they are not at fault, which may increase the price they charge to complete a task. Due to the potential increased cost associated with contracting real options, managers must conduct a detailed analysis of which potential options may be exercised in the future prior to signing contracts with vendors.

Like all financial forecasting, IRM relies on historical data to make predictions. Predictions that incorporate current information in their analysis rather than relying purely on historical trends can provide more insight to decision-makers. For example, sophisticated meteorologists create weather predictions from multiple sources: Doppler radar, satellites, radiosondes (i.e., weather balloons), and automated surface-observing systems observe the current weather conditions (National Oceanic and Atmospheric Administration [NOAA], 2017). Using numerical weather prediction, the data from the various sources is run in models based on known historical patterns for the region (NOAA, n.d.). For the meteorologist, knowing the current conditions is just as important as knowing the historical models (NOAA, n.d.). Similarly, if the project analyst had the ability to incorporate relevant information that is current to the minute (or to the required fidelity), the models would provide even more accurate information. Outsourcing, reducing manning, and maintaining the current structure all have statistics that could be used in simulations because of similar projects with historical data. Although this drawback is not unique to the IRM methodology, leaders should be aware of this flaw in any financial prediction.

Finally, the DoD does not currently incentivize PMs reaping the positive benefits of risk. The risk framework within DoD acquisitions is designed to minimize cost and schedule overruns during a project. The structure of DoD contracts does not encourage increased capability or performance from vendors or the project as a whole. Where a for profit business may invest in an endeavor that may fail, they do so because they believe the upside reward is greater than the potential cost of failure. For instance, if a design objective for an aircraft is to reach 250 knots and the design threshold is 200 knots, the budget will be for the threshold versus the objective. There will not be enough funding for the program to reach the objective unless the PM is able to reallocate resources internally. The acquisitions process looks at the cost to reach the objective rather the value of the objective. For-profit companies reward for performance, which is measured by revenue. The DoD’s unspoken surrogate for revenue is cost savings, which promotes a different value than increasing the value of a project. Conversely, DoD acquisitions proceed only once the negative consequences are mitigated to the greatest extent possible. The upside risk is of minimal importance to the PMs; the program simply needs to be completed on time and on budget. Although it is still vital to examine how potential projects fit into the overall collection of acquisitions and current assets in the DoD, the contract structure limits some of the portfolio optimization aspects of IRM.

Comparison of Key Attributes

Choosing a methodology should depend on the nature of the project under consideration, specifically, the commitment needed from the organization, the organization's desire to align strategic goals with the project, the predictive capability of the methodology, the flexibility required, and the time available. While others in the organization need to understand the concepts to comprehend status reports, EVM only needs the management team to track the cost and schedule of the project compared to the baseline as there is no determined goal alignment with the organization. While the CPI and SPI can help estimate the final cost and schedule, there is no true predictive ability associated with EVM since the assumption is that the schedule will proceed according to the baseline, regardless of previous performance volatility. Adherence to the baseline is essential in EVM, and changing requirements can drastically alter a baseline, reducing the effectiveness of the methodology. Setting up, monitoring, and reporting the cost/schedule performance of each work package within the WBS can be a time-consuming and expensive task for an AI project with its many unknowable components and capabilities a priori.

KVA needs only the KVA analyst and the process owner, as the SME, to determine the value of a process or component output, supporting the need to align the project with an organization's productivity goals. Using this analysis, they can model the current baseline as-is process ROK and compare it with the proposed to-be process model ROK, thus offering a simple prediction of the improvement between the models. Since KVA can be used with any language of description that defines the process outputs in common units, analysts can choose whichever method is most beneficial for the particular system in question, providing flexibility. This analysis can be completed quickly, potentially providing a rough-cut assessment within a few days. IRM requires the organizational leadership, portfolio and project managers, and the analyst to determine how a project fits within an organization's portfolio, the present value (PV) of the project, and potential real options. By analyzing and simulating various scenarios, IRM provides a prediction of a project's likely performance, which allows managers to build in flexibility via real options at the appropriate locations within the project. Assuming the data necessary for the analysis is available, the process can be completed in a relatively quick manner.

Methodologies in AI Acquisition

As previously discussed, the methodologies all have strengths and weaknesses, making them more suitable in certain applications than others. The biggest challenge in using EVM when acquiring AI is the iterative nature of software development. EVM needs clearly stated, detailed requirements for intermediate steps to be most effective. While the outputs of software programs are defined well, the steps required to build the software are not, leading to problems when developing cost and schedule estimates. If the software is not complex or consists of known processes, EVM can sufficiently monitor the progress. Integrating software and hardware is also complicated with EVM since there are numerous pieces of the program that must be combined to meet the goals, resulting in additional debugging and recoding. EVM is more efficient when used to manage the physical creation of systems or infrastructure. It can monitor the cost/schedule progress of software work packages but is not as useful at estimating the **value** of those programs.

KVA can provide an objective, ratio-scale measure of value and cost for each core process and its subprocesses or components within any IS system. Using the two parameters, managers can then analyze productivity ratios information, such as ROK and ROI, to determine the efficiency of a process compared to the resources used to achieve the output. This can help the manager decide how to use resources to update systems or estimate the future value of a system being acquired. Combining the KVA results with IRM allows managers to iterate the value of system real options analysis through simulation and other techniques. IRM can also quantify risks and forecast performance probabilities for measures of the potential success for programs and components of programs using historical data. It is a tool to assist with the investment strategy, making it useful when acquiring all types of AI. However, it is not designed to help manage the actual acquisition of an AI program or determine how to meet its detailed requirements.

Summary

Examining the benefits and challenges of the proposed methodologies demonstrates the scope, capabilities, and limitations of various AI systems. It also helps inform in which areas and phases of the Defense Acquisition System life cycle it may be appropriate to include the methodologies or components of the methodologies within the system. The next section provides recommendations based on these findings.

Conclusion

The main research question of this study was, simply, how can certain advanced analytical decision-making methodologies be used in the acquisition life cycle to complement existing methods to ensure a successful acquisition of AI technologies?

As discussed, EVM remains the only program management methodology required by the U.S. government for all DoD acquisition programs with a contract value exceeding $20 million. Regardless of this requirement, EVM is a methodology that provides a structured approach to the acquisition of IT via program management processes that can help ensure an acquisition program stays on schedule and within budgeted cost estimates. However, there are significant limitations when using EVM for AI acquisitions, the major weakness being that it was not designed for managing AI acquisitions that follow a very iterative and highly volatile pathway. Organic AI acquisitions require a high level of flexibility to deal with the unknowns that arise during the development process as well as value adding possibilities not in the original plan. In addition, EVM does not provide a common unit of value metric to enable standard productivity metrics, such as ROI. When value is inferred by how consistent a program is with original baseline cost and schedule estimates, the performance of the program may be sacrificed in terms of the quality of the outputs when planned program activities become iterative, as in the development of many AI programs. If an AI acquisition program is trending toward cost and schedule overruns, but the resulting value added of the modifications to the original requirements provides disproportionate increases in value, EVM is not designed to recognize this increase in value.

To remedy these shortcomings of EVM in AI acquisitions, the methodology should be combined with KVA and IRM, which can be useful during the requirements and monitoring phases of EVM 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 before acquisition of the an AI system. A future process model that estimates the value added of the incorporation of the AI can also set expectations that can be measured against the baseline model after the AI has been acquired. IRM can be used to forecast the value of strategic real options flexibility that an acquired AI may provide so that leadership can select the options that best fit their desired goals for the AI in defense core processes.

Because it provides an objective, quantifiable measure of value in common units, KVA should also be used in AI acquisitions to allow decision-makers to better understand and compare different strategic options based on their value and the cost. Obtaining a return on investment of AI systems can only be done when using KVA to determine the value embedded in the system. This information provides insight to PMs as well as a more complete perspective regarding the performance of both the current and the to-be systems.

Likewise, using IRM is recommended when acquiring AI through the Defense Acquisition System. Applying dynamic and stochastic uncertainty and risk-based modeling techniques to predict likely and probabilistic outcomes can improve the risk estimates associated with the components and subcomponents of a program, in terms of their potential cost overruns, value variabilities, and schedule delays. Analyzing various real options within the context of the models' outputs will help PMs make the most advantageous choices when determining a program's future.

PMs should use EVM only in the Engineering and Manufacturing Development (EMD) phase, as is currently done. That said, EVM will work best in hardware manufacturing solutions with technology that is fully mature prior to the program starting. Since many AI acquisition programs consist of advancing the current technology and developing new software solutions to meet requirements, EVM is not perfectly suited for AI development. Nevertheless, PMs can use various agile EVM techniques to complete projects on cost/schedule baselines provided the appropriate steps are taken when establishing the baseline. Requirements must be broken into small, easily definable tasks with suitable risk and uncertainty factors accounted for within the schedule. Other methodologies, such as KVA and IRM, should be used with EVM to ensure these factors are based on defendable metrics rather than simply guessing how much additional time, money, and value may be necessary to complete complex tasks.

During the Materiel Solution Analysis (MSA) phase, KVA and IRM will help determine the value of the different options considered in the analysis of alternative (AoA) process. KVA can objectively measure the value of the current, as-is system and the potential to-be systems under consideration. Then IRM can use additional factors to value the alternatives in terms of their relative parameter values such as cost, value, complexity, timeline. As the chosen solutions mature during the TMRR phase, an updated KVA analysis will reassess initial estimates and provide a projected ROI that can be incorporated in an IRM risk and real options analysis for the AI solution prior to entering the EMD phase as appropriate.

Limitations and Future Research

This research examined whether the various methodologies—EVM, KVA, and IRM—could be used within the Defense Acquisition System to improve the acquisition of AI. Future research should examine how these methodologies may interact with or improve other components of the acquisition system. This includes the JCIDS and PPBE components as individual processes and the interaction of JCIDS, PPBE, and the Defense Acquisition System as a whole. Certain methodologies, specifically IRM, may be more beneficial when used throughout the entire acquisition process instead of within a portion of the system. Additionally, future research could examine how these different methods may be used in the acquisition of products outside the AI or IT realm.

The research conducted looked at AI as a whole and not specific types of AI. Future studies should examine if acquisition methods, strategies, and methodologies should change based on the category of AI being acquired. This is of specific interest when considering artificial intelligence and its subsets. Machine learning, intelligence with a specific focus or field of expertise, and general or universal intelligence would likely have different methods used in the acquisition process based on their complexity, complicated nature, undeveloped technology, and level of risk.

The applicability of these methodologies within commercial acquisition of AI is another area of potential research. This research focused exclusively on the application of the respective techniques within the DoD acquisition process. However, commercial entities also struggle when acquiring complex or complicated AI and IT systems, particularly when the systems operate at the enterprise level. Further research may indicate if these same methodologies could provide value to decision-makers in the private sector during the creation, adoption, or customization of commercial AI. As noted in the literature, the hype cycle for AI and automation is on the rise and the demand to procure such technologies is as relevant for the commercial sector as it is for DoD. Furthermore, the recent pandemic caused by the Coronavirus Disease 2019 (COVID-19) has forced a permanent shift in society toward an increased trend toward a permanent remote workforce. As these trends are likely to continue in the foreseeable future, an increased in automation tools will be required to support this workforce. These trends could be explored for their implications as part of the Fourth Industrial Revolution and Industry 4.0.

Finally, this research examined only the most promising methodologies out of numerous different possibilities. Future research could examine other program management tools, management philosophies, analytic tools, or other methodologies and their benefit when acquiring AI. While the examined methodologies were chosen because they would likely benefit the process and support improvements in EVM, other systems may be more appropriate in certain phases or may offer additional benefits not seen in this research.

References

Apttus. (2015). *Success with PayPal*. https://solutions.apttus.com/rs/902-TNK-191/images/Apttus%20Case%20Study%20paypal.pdf

Apttus. (2017, September 20). *Apttus contract management: Using applied AI in contract management* [Video]. YouTube. https://www.youtube.com/watch?V=vthnaq2505y

Armstrong, C. S. (2015). *The quest for Achilles' Shield: Is the American military's fetish with technology harming mission accomplishment?* Joint Forces Staff College.

Ashby, W. R. (1956). *An introduction to cybernetics.* Chapman and Hall.

Baig, M. A. (2019, March 10). Gerasimov doctrine and modern hybrid warfare. *Daily Times.* https://dailytimes.com.pk/295075/gerasimov-doctrine-and-modern-hybrid-war/

Bartels, A. (2019). *The Forrester Wave™: Contract life cycle management for all contracts, Q1 2019*. Forrester Research.

Bech, H. (2019, September 26). *Silicon Valley's rule number one: Fake it till you make it*. https://tbkconsult.com/silicon-valleys-rule-number-one-fake-it-to-you-make-it/

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.

Bhatt, U. S., Newman, D. E., Carreras, B. A., & Dobson I. (2005). Understanding the effect of risk aversion on risk. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 64b. https://doi.org/10.1109/HICSS.2005.649

Bostron, N. (2014). *Superintelligence: Paths, dangers, strategies.* Oxford University Press.

Brandimarte, P. (2014). *Handbook in Monte Carlo simulation: Applications in financial engineering, risk management, and economics*. John Wiley & Sons.

Burkhalter, B. (1963). Applying artificial-intelligence to the pattern-cutters problem. *Operations Research*, *11*(1), B39.

Burton, J., & Soare, S. (2019). *Understanding the strategic implications of the weaponization of artificial intelligence* [Paper presentation]. 2019 11th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia.

Burton, S. (2018, February). The case for plain-language contracts. *Harvard Business Review*. https://hbr.org/2018/01/the-case-for-plain-language-contracts

Buss, T. F., & Cooke, D. (2005, May). *Performance measurement in defense acquisitions: A case study of the Navy* [Paper presentation]. Acquisition Research Symposium, Naval Postgraduate School, Monterey, CA.

Capra, F. (1996). *The web of life: A new scientific understanding of living systems.* Anchor Books.

Carlton, B., Ellis, K., Jones, J., & Schofield, B. (2019). *Marine Corps transport corrective maintenance process optimization and knowledge value analysis*. Unpublished manuscript.

Cassidy Law. (2018, September 21). *The dread-inducing review of a federal government RFP or contract*. http://cassidylawpllc.com/review-of-a-federal-government-rfp-or-contract/

Chaplain, C. (2017). *Space acquisitions: DoD continues to face challenges of delayed delivery of critical space capabilities and fragmented leadership* (GAO-17-619T). Government Accountability Office.

Chaplain, C. (2019). *DoD faces significant challenges as it seeks to address threats and accelerate space programs* (GAO-19-482T). Government Accountability Office.

Char, J., & Bitzinger, R. (2016, October 5). Reshaping the People's Liberation Army since the 18th Party Congress: Politics, policymaking, and professionalism. *Journal of Strategic Studies*, *39*(5–6), 599–607. https://doi.org/10.1080/01402390.2016.1235037

Christensen, D. (1994). An analysis of cost overruns on defense acquisition contracts. *Readings in Program Control*. https://ntrs.nasa.gov/search.jsp?R=19950014332

Christensen, D. S. (1998). The costs and benefits of the Earned Value Management process. *Acquisition Research Quarterly*, *5*(4), 373–386. https://www.researchgate.net/publication/251811135_The_Costs_and_Benefits_of_the_Earned_Value_Management_Process

Congressional Research Service. (2018). *The global research and development landscape and implications for the Department of Defense*.

Cook, R. I. (2000). *How complex systems fail*. Massachusetts Institute of Technology. https://web.mit.edu/2.75/resources/random/How%20Complex%20Systems%20Fail.pdf

Cooper, D. E. (2005). *Contract management: Opportunities to improve surveillance on Department of Defense service contracts* (GAO-05-274).

Creviston, D. O. (2020). Transforming DoD for agile multidomain command and control. *Joint Force Quarterly: JFQ*, (97), 83–90.

Crosby, C. (2020). Operationalizing artificial intelligence for algorithmic warfare. *Military Review*, *100*(4), 42–51.

Cross, J. (2019, May 9). Department of Defense emerging technology strategy: A venture capital perspective. *Proceedings of the 16th Annual Acquisition Research Symposium*. https://calhoun.nps.edu/bitstream/handle/10945/63024/SYM-AM-19-040.pdf?Sequence=1&isallowed=y

Cummins, T. (2019, October). *The cost of a contract*. https://blog.iaccm.com/commitment-matters-tim-cummins-blog/the-cost-of-a-contract

Daley, S. (2019, September 24). *19 examples of artificial intelligence shaking up business as usual*. https://builtin.com/artificial-intelligence/examples-ai-in-industry

Dougherty, J., & Jay, M. (2018). Russia tries to get smart about artificial intelligence. *The Wilson Quarterly.* https://www.wilsonquarterly.com/quarterly/living-with-artificial-intelligence/russia-tries-to-get-smart-about-artificial-intelligence/

Damodaran, A. (2000). The promise of real options. *Journal of Applied Corporate Finance*, *13*(2), 29–44. https://doi.org/10.1111/j.1745-6622.2000.tb00052.x

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

DARPA. (2019). *AI next campaign*. https://www.darpa.mil/work-with-us/ai-next-campaign

Defense Procurement and Acquisition Policy. (2012, April). *Defense contingency contracting handbook*. Office of the Under Secretary of Defense for Acquisition. https://www.acq.osd.mil/dpap/ccap/cc/corhb/Files/DCCOR_Handbook_2012.pdf

Defense Science Board. (2018, February). *Design and acquisition of software for defense systems*. https://apps.dtic.mil/dtic/tr/fulltext/u2/1048883.pdf

Denning, P. (2019, September). *Harnessing artificial intelligence*. Naval Postgraduate School.

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

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

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

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

Department of Defense. (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/

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

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

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

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

Department of Defense. (2019a, January 18). *Earned value management implementation guide (EVMIG)*.

Department of Defense. (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

Department of Homeland Security. (2019). *CPARS artificial intelligence commercial solutions opening pilot program general solicitation* (CSOP-HQ-GS-000001 Amendment 1).

DiNapoli, T.J. (2019). *DoD's use of other transactions for prototype projects has increased* (GAO-20-84). Government Accountability Office.

Dixit, A. K., & Pindyck, R. S. (1995). The options approach to capital investment. *Harvard Business Review*, *73*(3). https://hbr.org/1995/05/the-options-approach-to-capital-investment

DOHS. (2019, August 9). *Artificial intelligence (AI) for past performance—CPARS*. https://gov.surf/opportunitydetails/Index?Id=CSOP-HQ-GS-000001&originalposteddate=08/05/2019%2005:00:01

Exec. Order 13859. (2019, February 11). *Maintaining American leadership in artificial intelligence*. https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence

Edwards, C., & Kaeding, N. (2015) Federal government cost overruns. *Tax & Budget Bulletin, 72*. https://object.cato.org/sites/cato.org/files/pubs/pdf/tbb-72.pdf

Electronic Industries Alliance. (1998). *Earned Value Management systems.* Electronic Industries Alliance.

Egel, D., Robinson, E., Cleveland, C. T., & Oates, C. J. (2020, January 27). AI and irregular warfare: An evolution, not a revolution. *War on the Rocks*: https://warontherocks.com/2019/10/ai-and-irregular-warfare-an-evolution-not-a-revolution/

Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework* (Summary Report AFOSR-3233). Stanford Research Institute.

Evron, Y. (2012, February). China's military procurement approach in the early 21st century and its operational implications. *Journal of Strategic Studies*, 63–93. https://doi.org/10.1080/01402390.2011.592004

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.

Federal Acquisition Regulation (FAR), 48 C.F.R. 1.102 (2019).

Feickert, A. (2005). *The Joint Tactical Radio System (JTRS) and the Army's future combat system (FCS): Issues for Congress* (CRS Report No. RL33161). https://digital.library.unt.edu/ark:/67531/metacrs7941/

Fischetti, M. P. (2018, January 2). Yes, it can be done: Expatiating defense acquisition. *Federal Times.* https://www.federaltimes.com/opinions/2018/01/02/yes-it-can-be-done-expediting-defense-acquisition/

Fleming, Q. W., & Koppelman, J. M. (2010). *Earned value project management* (4th ed.). Project Management Institute, Inc. https://app.knovel.com/hotlink/toc/id:kpEVPME001/earned-value-project/earned-value-project

Francis, P. L. (2006). *Defense acquisition: Restructured JTRS program reduces risk, but significant challenges remain* (GAO-06-955). Government Accountability Office.

Francis, P. L. (2008). *Defense of Defense needs framework for balancing investments in tactical radios* (GAO-08-877). Government Accountability Office.

Friedman, N. (2013). *This truck saved my life*. Joint Program Office Mine-Resistant Ambush-Protected Vehicles.

Fuller, J. B., & Masko, J. (2019, September). *Theranos: The unicorn that wasn't*. Harvard Business School. https://www.hbs.edu/faculty/Pages/item.aspx?Num=55762

Future of Life Institute. (2020, August 15). *AI open letter*. https://futureoflife.org/ai-open-letter/?cn-reloaded=1

Gallagher, S. (2012, June 18). *How to blow $6 billion on a tech project*. Ars Technica. https://arstechnica.com/information-technology/2012/06/how-to-blow-6-billion-on-a-tech-project/

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

Geist, E. (2016). It's already too late to stop the AI arms race—We must manage it instead. *Bulletin of the Atomic Scientists*, *72*(5), 318–21.

Gerding, E. F. (2013). Contract as a pattern language. *Washington Law Review*, *88*(4), 1323–1356. https://digitalcommons.law.uw.edu/wlr/vol88/iss4/6/

Gerenser, B. (2019). DoD launches travel system prototype. *Defense Transportation Journal*, *75*(2), 10–11.

Giachetti, R. E. (2015). Evaluation of the DoDAF meta-model's support of systems engineering. *Procedia Computer Science*, *61*, 254–260.

Giannetti, W. (2020). Quiet giant: The TITAN cloud and the future of DoD artificial intelligence. *Air & Space Power Journal*, *34*(1), 54–58.

Gibbs, S. (2014, October 27). Elon Musk: Artificial intelligence is our biggest existential threat. *The Guardian*. https://www.theguardian.com/technology/2014/oct/27/elon-musk-artificial-intelligence-ai-biggest-existential-threat

Golden, P. E. (2020). DoD's artificial intelligence problem: Where to begin. *The Army Lawyer*, 2, 76–85.

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

Good, A. (2019, November 8). *Intelligent failure assessment*. Fail Forward. https://static1.squarespace.com/static/583382786b8f5b1d0c788b9e/t/58af22a6e4fcb57eb0f01028/1487872679331/Fail+Forward+Intelligent+Failure+Assessment.pdf

Gourley, S. (2012, November 28). *Rapid Fielding Initiative: A decade of providing urgently needed gear*. Defense Media Network. https://www.defensemedianetwork.com/stories/rapid-fielding-initiative/

Government Accountability Office. (2017). *Military acquisitions: DoD is taking steps to address challenges faced by certain companies* (GAO-17-644).

Government Accountability Office. (2019a, May 21). *2019 annual report: Additional opportunities to reduce fragmentation, overlap, and duplication and achieve billions in financial benefits* (GAO-19-285SP). https://www.gao.gov/reports/GAO-19-285SP/

Government Accountability Office. (2019b). *DoD's use of other transactions for prototype projects has increased* (GAO-20-84). Government Accountability Office.

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/

Gubrud, M. (2014). Stopping killer robots. *Bulletin of the Atomic Scientists*, *70*(1), 32–42.

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

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On past, present, and future of AI. *California Management Review*, *61*(4), 5–14.

Hagan, G. (2009, November). *Glossary of defense acquisition acronyms and terms*. Defense Acquisition University. http://www.acqnotes.com/Attachments/DAU%20-%2013th_Edition_Glossary.pdf

Hagel, C. (2014a, November 15). *The defense innovation initiative*. Defense.gov Archive. https://archive.defense.gov/pubs/OSD013411-14.pdf

Hagel, C. (2014b, November 15). *Reagan National Defense Forum keynote* [Speech]. https://www.defense.gov/Newsroom/Speeches/Speech/Article/606635/

Heller, C. H. (2019). Near-term applications of artificial intelligence: Implementation opportunities from modern business practices. *Naval War College Review*, *72*(4), 79–105.

Hillner, E. P. (2019, May). *The Third Offset Strategy and the Army modernization priorities*. Center for Army Lessons Learned. https://usacac.army.mil/sites/default/files/publications/17855.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 presented at the 16th Acquisition Research Symposium, Monterey, CA.

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

Hutton, J. (2011). *Defense contract management agency: Amid ongoing efforts to rebuild capacity, several factors present challenges in meeting its missions* (GAO-12-83). Government Accountability Office.

Icertis. (n.d.-a). Contract management software. Retrieved October 11, 2019, from https://www.icertis.com/contract-management-software/

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/

Ilachinski, A. (2017, October). *AI, robots, and swarms: Issues, questions, and recommended studies*. Center for Naval Analyses. https://apps.dtic.mil/dtic/tr/fulltext/u2/1041749.pdf

Inhofe, J., & Reed, J. (2019). FY 2020 National Defense Authorization Act. Washington, DC: Senate Armed Services Committee.

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

Institute for the Study of War. (2019). *Provincial reconstruction teams*.

Jones, R. (2019). *Procurement and contracting: Overview and primary considerations*. Naval Postgraduate School.

Jones, R., & Housel, T. (2018, May 9). *Extending an econophysics value model for early developmental program performance prediction and assessment*. Paper presented at the 15th Acquisition Research Symposium, Monterey, CA.

Kania, E. B. (2019, August 27). *In military–civil fusion, China is learning lessons from the United States and starting to innovate*. Real Clear Defense. https://www.realcleardefense.com/articles/2019/08/27/in_military-civil_fusion_china_is_learning_lessons_from_the_united_states_and_starting_to_innovate_114699.html

Kaplan, A. M., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62 (1), 15–25.

Kashin, V., & Raska, M. (2017). *Countering the Third Offset Strategy: Russian perspectives, responses, and challenges*. S. Rajaratnam School of International Studies.

Keirsey, D., Mitchell, J., Bullock, B., Nussmeier, T., & Tseng, D. Y. (1985). Autonomous vehicle control using AI techniques. *IEEE Transactions on Software Engineering*, *11*(9), 986–992.

Kendall, F. (2011, October 13). *Letter to the Honorable Howard K. "Buck" McKeon*. https://www.govexec.com/pdfs/101411bb1.pdf

Kendall, F. (2017). *Getting defense acquisition right*. DAU Press. https://DoD.defense.gov/Portals/1/Documents/pubs/Getting-Acquisition-Right-Jan2017.pdf

Kendall, F. (2018, March 20). *Five myths about Pentagon weapons programs*. Defense One. https://www.defenseone.com/ideas/2018/03/five-myths-about-pentagon-weapons-programs/146803/

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

Kissinger, H. A. (2019, August 30). How the Enlightenment ends. *The Atlantic.* https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/

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/

Kovacs, G., & Kot, S. (2016). New logistics and production trends as the effect of global economy changes. *Polish Journal Management Studies*, *14*(1), 122.

Kurzweil, R. (2000). *The age of spiritual machines: When computers exceed human intelligence*. Penguin Books.

Kwak, Y. H., & Anbari, F. T. (2011). History, practices, and future of earned value management in government: Perspectives from NASA. *Project Management Journal*, *43*(1), 77–90. https://doi.org/10.1002/pmj.20272

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

Lee, J. (2013). Industry 4.0 in big data environment. *German Harting Magazine, Technology Newsletter*, 26.

Lee, U. (2014). *Improving the parametric method of cost estimating relationships of naval ships* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/92134

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

Levine, P., & Greenwalt, B. (2019, April 4). *What the 809 panel didn't quite get right*. Breaking Defense. https://breakingdefense.com/2019/04/what-the-809-panel-didnt-quite-get-right-greenwalt-levine/

Li, L., Zheng, N., & Wang, F. (2019). On the crossroad of artificial intelligence: A revisit to Alan Turing and Norbert Wiener. *IEEE Transactions on Cybernetics*, *49*(10), 3618–3626.

Licklider, J. C. R. (1960). Man–computer symbiosis. *IRE Transactions on Human Factors in Electronics*, *4*(11).

Ludwigson, J. (2019). *DoD space acquisitions: Including users early and often in software development could benefit programs* (GAO-19-136). Government Accountability Office.

Mackin, M. (2015). *Littoral Combat Ship: Knowledge of survivability and lethality capabilities needed prior to making major funding decisions* (GAO-16-201). Government Accountability Office.

Maucione, S. (2019a, April 26). *DoD doesn't want to end up like Theranos, so it's being cautious about "fail fast."* Federal News Network. https://federalnewsnetwork.com/defense-main/2019/04/DOD-doesnt-want-to-end-up-like-theranos-so-its-being-cautious-about-fail-fast/

Maucione, S. (2019b, October 30). *Special report: Failure is an option for DoD's experimental agency, but how much? Federal News Network.* https://federalnewsnetwork.com/defense-main/2019/10/special-report-failure-is-an-option-for-DoDs-experimental-agency-but-how-much/

McCain, J. (2015, November 15). It's time to upgrade the defense department. *War on the Rocks*. https://warontherocks.com/2015/11/its-time-to-upgrade-the-defense-department/

McCusker, E. (2019). *Department of Defense fiscal year 2020*. Under Secretary of the Department of Defense Comptroller.

Mori, S. (2018). U.S. defense innovation and artificial intelligence. *Asia-Pacific Review*, *25*(2), 16–44.

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 presented at Naval Acquisitions Research Conference, Monterey, CA.

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

Mun & Anderson. (2021*).* Technology trust: system information impact on autonomous systems adoption in high-risk applications. *Defense Acquisition Research Journal, January 2021 (Vol. 28 No. 1:2-39).* DOI: 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*, Volume 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

National Defense Transportation Association. (2020, January 28). *DoD launches travel system prototype*. https://www.ndtahq.com/dod-launches-travel-system-prototype/

National Oceanic and Atmospheric Administration. (n.d.). *Numerical weather prediction*. Retrieved April 17, 2019, from https://www.ncdc.noaa.gov/data- access/model-data/model-datasets/numerical-weather-prediction

National Oceanic and Atmospheric Administration. (2017, August 14). *6 tools our meteorologists use to forecast the weather*. https://www.noaa.gov/stories/6-tools-our-meteorologists-use-to-forecast-weather

National Science Foundation. (2018). *National patterns of R&D resources: 2016–2017 data update.*

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

Nielsen, C. B., Larsen, P. G., Fitzgerald, J., Woodcock, J., & Peleska, J. (2015). Systems of systems engineering: Basic concepts, model-based techniques, and research directions. *ACM Computing Surveys*, *48*(2), 18:1–18:41.

Oakley, S. S. (2020). *Defense acquisitions annual assessment: Drive to deliver capabilities faster increases importance of program knowledge and consistent data for oversight* (GAO-20-439). Government Accountability Office.

Office of Small Business Programs. (2019, November 5). *Rapid Innovation Fund*. https://business.defense.gov/Programs/RIF/

Office of the Under Secretary of Defense for Acquisition & Sustainment. (2019a, November 1). *Ellen M. Lord*. https://www.acq.osd.mil/bio_lord.html

Office of the Under Secretary of Defense for Acquisition & Sustainment. (2019b, November 1). *OUSD(A&S) strategy road map*. https://www.acq.osd.mil/fo/docs/as-roadmap.pdf

Office of the Under Secretary of Defense for Research and Engineering. (2019, December 2). *USD(R&E) modernization priorities*. https://www.cto.mil/modernization-priorities/

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/

Payne, K. (2018). *Strategy, evolution, and war: From apes to artificial intelligence*. Georgetown University Press.

Pellerin, C. (2015, November 8). *Work: Human–machine teaming represents defense technology future*. Department of Defense. https://www.defense.gov/Explore/News/Article/Article/628154/workhu/

People's Republic of China Ministry of Science and Technology. (2018). *National High-Tech Program (863 Program)*. http://www.most.gov.cn/eng/programmes1/

Pickens, A. H., & Alvarado, D. J. (2018). *Other transaction agreements: An analysis of the oracle decision and its potential impact on the use of OTAs*. American Bar Association.

Porche, I. R., Mckay, S., Mckernan, M., Button, R. W., Murphy, B. Giglio, K., & Axelband, E. (2012). *Rapid acquisition and fielding for information assurance and cyber security in the Navy*. RAND Corporation.

Powner, D. A. (2009). *Information technology: Agencies need to approve the implementation and use of earned value management techniques to help manage major system acquisitions* (GAO-10-2). Government Accountability Office.

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

Ramsey, L. (2019, June 28). Theranos founder Elizabeth Holmes faces jail time for fraud charges. Her trial is set to begin in summer 2020. *Business Insider.* https://www.businessinsider.com/theranos-founder-elizabeth-holmes-president-sunny-balwani-trial-date-2019-6

Rashid, M., Anwar, M. W., & Khan, A. M. (2015). Toward the tools selection in model based system engineering for embedded systems—A systematic literature review. *The Journal of Systems and Software*, *106*, 150–163.

Ramos, A. L., Ferreira, J. V., & Barceló, J. (2012). Model-based systems engineering: An emerging approach for modern systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *42*(1), 101–111.

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

Rios, C. (2005). *Return on investment analysis of information warfare systems* (Master's thesis, Naval Postgraduate School). https://calhoun.nps.edu/handle/10945/2068

Rios, C., Housel, J., & Mun, J. (2006). *Integrated portfolio analysis on investment and real options analysis of intelligence information systems (Cryptologic Carry on Program)*. Paper presented at the Acquisition Research Symposium, Monterey, CA. https://calhoun.nps.edu/handle/10945/33143

Rios, C., Housel, J., & Mun, J. (2015). Case study 11: Real options and KVA in military strategy at the United States Navy. In J. Mun (Ed.), *Case studies in certified quantitative risk management (CQRM)* (pp. 145–154). Thomson-Shore and ROV Press.

Roberts, S. (2017, February 11). Leonard Perroots, general who defused nuclear crisis with Soviets, dies at 83. *New York Times.* https://www.nytimes.com/2017/02/10/us/leonard-perroots-dead.html

Roosevelt, A. (2012, October 2). Joint Tactical Networking Center opens, JPEO JTRS officially closed. *Defense Daily.* https://www.defensedaily.com/joint-tactical-networking-center-opens-jpeojtrsofficially-closed/navy-usmc/

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

Sargent, J. F., Gallo, M. E., & Schwartz, M. (2018, November 8). *The global research and development landscape and implications for the Department of Defense*. Congressional Research Service. https://fas.org/sgp/crs/natsec/R45403.pdf

Sayler, K. M. (2020). *Artificial intelligence and national security* (CRS Report No. R45178). https://fas.org/sgp/crs/natsec/R45178.pdf

Schaeffer, C. (2016, July 12). *Agile versus Waterfall for CRM Implementation success*. Customer Think. https://customerthink.com/agile-versus-waterfall-for-crm-implementation-success/

Schlimmer, S., & Brennan, B. (2018, July 4). *For vendors, DoD's CUI requirements more than an exercise*. Federal News Network. https://federalnewsnetwork.com/commentary/2018/07/for-vendors-DoDs-cui-requirements-are-more-than-an-exercise/

Schwab, K. (2015, December 12). The Fourth Industrial Revolution: What it means and how to respond. *Foreign Affairs*. https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution

Schwab, K. (2017). *The Fourth Industrial Revolution*. Crown Business.

Schwartz, M., Sargent, J. F., & Mann, C. T. (2018, July 2). *Defense acquisitions: How and where DoD spends its contracting dollars*. https://www.hsdl.org/?view&did=812910

Scott, W. A. (2007). *Request for and use of emergency supplemental funds for the Rapid Fielding Initiative*. Inspector General United States Department of Defense.

Section 809 Panel. (2018). *Report of the advisory panel of streamlining and codifying acquisition regulations* (Vol. 1).

Section 809 Panel. (2019b). *A road map to the Section 809 Panel reports*.

Section 809 Panel. (2019a). *Report of the advisory panel on streamlining and codifying acquisition regulations* (Vol. 3).

Selten, R. (1991). *Game equilibrium models IV social and political interaction*. Springer.

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

Sharkey, N. (2007, August 17). Robot wars are a reality. *The Guardian*. https://www.theguardian.com/commentisfree/2007/aug/18/comment.military

Shaw, M. (2019, October 15). *Why Google is the best search engine (and why businesses should care)*. Tower Marketing. https://www.towermarketing.net/blog/google-best-search-engine/

Shin, Y., Sim, S., & Lee, J. (2017). Model-based integration of test and evaluation process and system safety process for development of safety-critical weapon systems. *Systems Engineering*, *20*(3), 257–279.

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

Simpson, J. (2018, March 12). *FY18 procurement action lead time (PALT) metric*. PEOSTRI. https://www.peostri.army.mil/palt-memo

Oppy, G., & Dowe, D. (2016, February 8). *The Turing Test*. In E. N. Zalta (ed.), *Stanford encyclopedia of philosophy.* https://plato.stanford.edu/entries/turing-test/

Sparkes, M. (2015, January 13). Top scientists call for caution over artificial intelligence. *The Telegraph (UK).*

Statistics Solutions. (n.d.). *Data levels of measurement*. Retrieved September 9, 2019, from https://www.statisticssolutions.com/data-levels-of-measurement/

Stewart, R. (2019, July 26). *AFCEA Shark Tank - Rapid Acquisition Facilitation Tool (R.A.F.T.)*. Federal Government Experts. https://www.federalgovernmentexperts.com/post/afcea-shark-tank-rapid-acquisition-facilitation-tool-r-a-f-t

Sullivan, M. (2009). *Rapid acquisition of MRAP vehicles*. Government Accountability Office.

Sullivan, M. (2015). *Defense Advanced Research Project Agency: Key factors drive transition of technologies, but better training and data dissemination can increase success* (GAO-16-5). Government Accountability Office.

Sullivan, M.J. (2018). *Defense science and technology: Actions needed to enhance use of laboratory initiated research authority* (GAO-19-64). Government Accountability Office.

Tarraf, D. C., Shelton, W., Parker, E., Alkire, B., Gehlhaus, D., Grana, J., Levedahl, A., Leveille, J., Mondschein, J., Ryseff, J., Wyne, A., Elinoff, D., Geist, E., Harris, B.N., Hui, E., Kenney, C., Newberry, S., Sachs, C., Schirmer, P., Schlang, D., … Warren, K. (2019). *The Department of Defense posture for artificial intelligence: Assessment and recommendations*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR4229.html

Tatum, D. (2018, June). *Contracting Officer's Representative (COR): An analysis of part-time and full-time COR roles, competency requirements and effectiveness*. https://calhoun.nps.edu/bitstream/handle/10945/59604/18Jun_Tatum_Denise.pdf?Sequence=1&isallowed=y

Taylor, T. (2019). Artificial intelligence in defence. *The RUSI Journal*, *164*(5-6), 72–81.

National Security Archive. (2019, February 26). *The Able Archer 83 Sourcebook*. https://nsarchive.gwu.edu/project/able-archer-83-sourcebook

Thomas, E., Christina, C., & Painter, G. (2019). *Leveraging technology to improve contract surveillance: Opportunity identification and analysis*. Naval Postgraduate School.

National Defense Authorization Act for Fiscal Year 2016, Pub. L. 114-92 (2015). https://www.congress.gov/114/plaws/publ92/PLAW-114publ92.pdf

Tsai, A. (2014). An empirical model of four processes for sharing organizational knowledge. *Online Information Review*, *38*(2), 305–320. https://doi.org/10.1108/OIR-03-2013-0059

U.S. House of Representatives. (2017). National Defense Authorization Act for fiscal year 2018. Washington, DC: House of Representatives.

Office of the Under Secretary of Defense. (2003). *Manager's guide to technology transition in an evolutionary acquisition environment*.

Under Secretary of Defense, Acquisition, Technology, and Logistics. (2016). *Performance of the defense acquisition system: 2016 annual report.* Department of Defense.

Vanhoucke, M. (2014). *Integrated project management and control*. Springer International Publishing.

Van Eck, N. J., & Waltman, L. (2020). *VOSviewer - Visualizing scientific landscapes*. https://www.vosviewer.com/

Van Schooten, D. (2018, August 16). *The Pentagon's contracting gurus mismanaged their own contracts: The Defense Contract Management Agency botched a $45-million project to help manage trillions of dollars in other contracts*. Project on Government Oversight. https://www.pogo.org/investigation/2018/08/pentagons-contracting-gurus-mismanaged-their-own-contracts/

Vincent, J. (2019, January 28). *The state of AI in 2019.* The Verge. https://www.theverge.com/2019/1/28/18197520/ai-artificial-intelligence-machine-learning-computational-science

Voelz, G. (2016). Catalysts of military innovation: A case study of defense biometrics. *Defense Acquisition Research*, *23*(2), 178–201.

Waikar, S. (2018, December 17). *What can we learn from the downfall of Theranos?* Stanford Graduate School of Business. https://www.gsb.stanford.edu/insights/what-can-we-learn-downfall-theranos

Walton, T. (2016). *Securing the Third Offset Strategy: Priorities for the next secretary of defense*. NDU Press. https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-82/jfq-82_6-15_Walton.pdf

Web of Science. (n.d.). Retrieved October 26, 2020, from http://apps.newisiknowledge.com/

Weisgerber, M. (2017, December 21). The Pentagon's new artificial intelligence is already hunting terrorists. *Defense One*. http://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine.* MIT Press.

White House. (2020). *National strategy for critical and emerging technologies.* https://www.whitehouse.gov/wp-content/uploads/2020/10/National-Strategy-for-CET.pdf

Wimmer, H., & Rada, R. (2018). Applying artificial intelligence to financial investing. In *The Encyclopedia of Information Science and Technology*. IGI Global.

Work, R. O. (2014, November 12). *Deputy secretary of defense speech at the CSIS Global Security Forum*. Department of Defense. https://www.defense.gov/Newsroom/Speeches/Speech/Article/606631/

Work, R. O. (2015, January 28). *The Third U.S. Offset Strategy and its implications for partners and allies*. Department of Defense. https://www.defense.gov/Newsroom/Speeches/Speech/Article/606641/the-third-us-offset-strategy-and-its-implications-for-partners-and-allies/

Work, R., & Schmidt, E. (2019, July 17). In search of ideas: The National Security Commission on Artificial Intelligence wants you. *War on the Rocks:* https://warontherocks.com/2019/07/in-search-of-ideas-the-national-security-commission-on-artificial-intelligence-wants-you/

Yao, M., Jia, M., & Zhou, A. (2018). *Applied artificial intelligence: A handbook for business leaders*. TOPBOTS Inc.

Young, J. J., Jr.. (2004, November 30). *Blueprint for the future, Naval Research and Acquisition Team 1999–2004 strategic plan* [Memorandum]. Department of Defense. https://www.secnav.navy.mil/rda/Pages/ViewPolicy.aspx

Yu, W., Chang, P., Yao, S., & Liu, S. (2009). KVAM: model for measuring knowledge management performance of engineering community of practice. *Construction Management and Economics*, *27*(8), 733–747. https://doi.org/10.1080/01446190903074978

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/

Zoroya, G. (2013, December 18). How the IED changed the U.S. military. *USA Today.* https://www.usatoday.com/story/news/nation/2013/12/18/ied-10-years-blast-wounds-amputations/3803017/

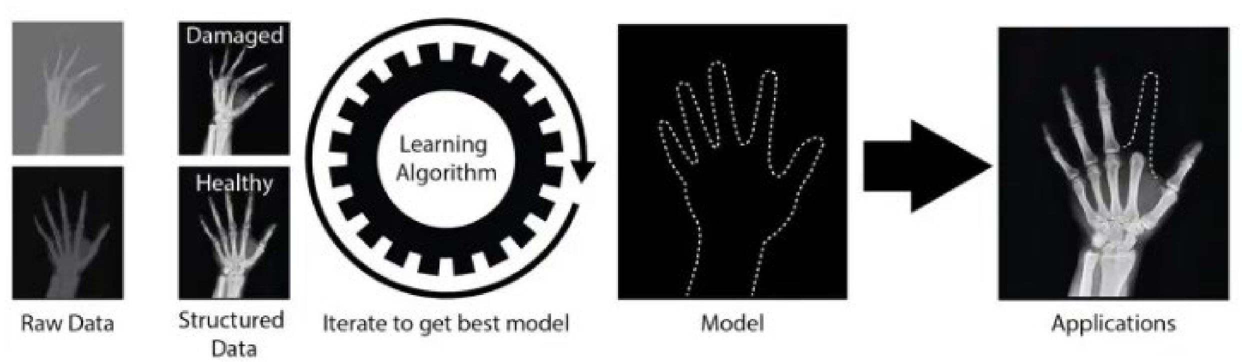
Appendix A: AI Capabilities

This Appendix provides amplifying information on the AI capabilities that were introduced in the third section of this report.

Machine Learning

Intelligence is the capacity to process a certain type of information, allowing a processor to solve problems that are of consequence (Gardner, 1993). Multiple types of intelligence have been proposed by psychologists beyond the classic understanding of a person's analytic intelligence quotient (IQ), which can sometimes measure only how well someone takes an IQ test instead of one's innate abilities. Howard Gardner proposed a theory of multiple intelligence, which suggests that the traditional psychometric views of intelligence are too narrow and should include more categories where certain processors, people in this instance, are stronger than others at making sense of different stimuli. These categories of intelligence include visual-spatial, linguistic-verbal, interpersonal, intrapersonal, logical-mathematical, musical, body-kinesthetic, and naturalistic (Gardner, 1993). An argument against this proposition would be that these categories merely represent learned and disciplined behaviors that someone developed in their life because of personality and circumstances. Regardless, both concepts of intelligence (classic and multiple) play well into what steps are taken to create a machine with artificial intelligence.

A computer can perform calculations based on the data provided and ultimately returns an a priori determined result. It can be programmed and coded to follow certain steps or algorithms repeatedly, and even alter its findings based on its own previously calculated results through some error correction algorithms. A combination of these two steps is the basic concept of machine learning. A computer system is first fed data structured in a way that the algorithm is programmed to recognize, derive patterns from the data, and make assumptions about any unstructured data provided subsequently (Greenfield, 2019). Figure 34 shows how this works in an x-ray learning algorithm.



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 34. AI training algorithm

***Source:* Greenfield (2019)**

Figure 34 simply illustrates the fundamental concept of machine learning, but the current research focuses on the different types of learning from the perspective of procurement. Sievo (2019), an AI procurement software company, provides the following are interpretations of different types of learning in the procurement algorithms.

Supervised Learning

An algorithm is taught the patterns using past data, and then detects them automatically in new data. Supervision comes in the form of correct answers that humans provide to train the algorithm to seek out patterns in data. This is commonly used within procurement areas such as spend classification (Sievo, 2019).

Unsupervised Learning

The algorithm is programmed to detect new and interesting patterns in completely new data. Without supervision, the algorithm is not expected to surface specific correct answers; instead, it looks for logical patterns within raw data. This is rarely used within critical procurement functions (Sievo, 2019).

Reinforcement Learning

The algorithm decides how to act in certain situations, and the behavior is rewarded or punished depending on the consequences. This is largely theoretical in the procurement context (Sievo, 2019).

Deep Learning

An advanced class of machine learning inspired by the human brain where artificial neural networks progressively improve their ability to perform a task. This is an emerging opportunity in procurement functions (Sievo, 2019).

Natural Language Processing

Anyone familiar with gadgets that are seemingly able to understand written or spoken words and act on them, such as translation apps or personal assistants like Amazon's Alexa, are already interacting with NLP-enabled AI. NLP incorporates algorithms that can interpret, transform, and generate human language in a way that is clearly understood by people (Sammalkorpi & Teppala, 2019). Speech soundwaves are transformed into computer code that has meaning to the algorithms. The code then converts that meaning back into a human understandable, accurate response that can be applied to normal human cognition. This process is accomplished through semantic parsing that maps the language of a passage to categorize each word and through machine learning, makes associations to denote not just the definition of the word, but its meaning in context (Raghaven & Mooney, 2013). Figure 35 illustrates this process of categorization and analysis in context for a portion of a procurement contract.

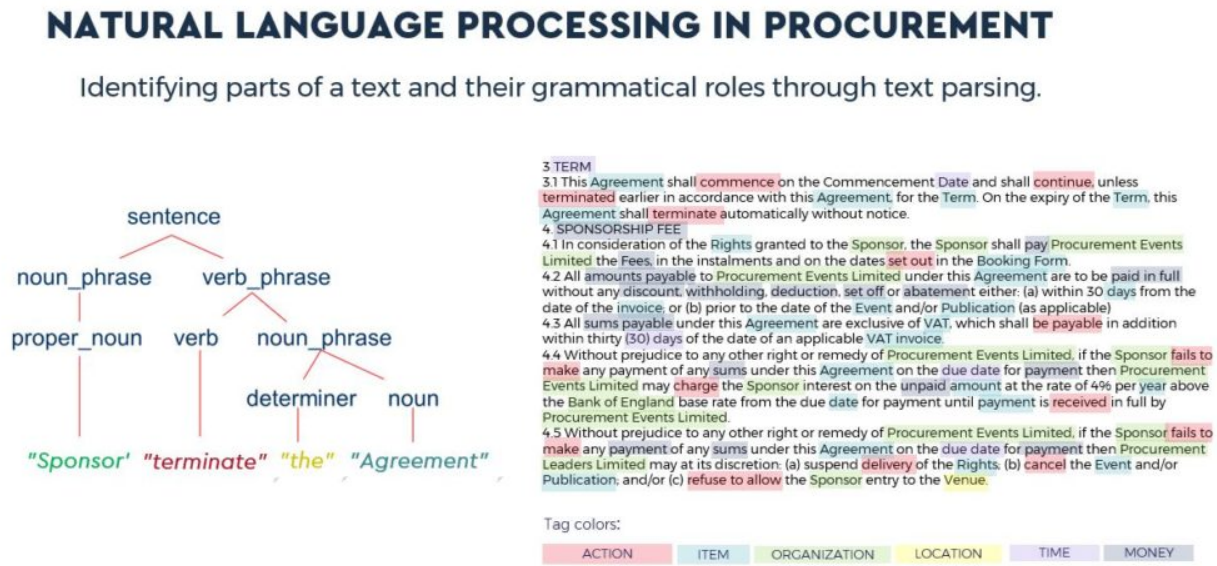


Figure 35. Semantic parsing in procurement

Source: **Sievo (2019)**

Robotic Process Automation

As discussed in the third section of this report, RPA is not AI; rather, it is an existing process enhanced by AI. By definition, RPA is "the application of technology that allows employees in a company to configure computer software or a robot to capture and interpret existing applications for processing a transaction, manipulating data, triggering responses, and communicating with other digital systems" (IRPA & AI, 2019). Robotic automation has many advantages when it can be applied appropriately because it is not limited by human constraints in terms of fatigue, morale, discipline, and survival needs. Unlike their human makers, robots also lack aspirations. Working harder will not result in promotions or more income, and even being permanently shut off bears no significance because robotic automation only mimics practical aspects of human intelligence and not the deeper nature of humanity itself (Zarkadakis, 2019). (Note, however, that ML does rely on an incentive mechanism for the machine to make decisions in terms of positive or negative responses.)

A possibility for future AI-enabled RPA is for a machine to learn how to manipulate the source of positive reinforcement completely out of following the rules needed to meet its goal. Things that survive, evolve to continue their survival and do so out of positive reinforcement from their environment, and the fact that they continue to act in a way deemed survivable. This should be considered in any future AI projects, and even more in the case why a human must always be involved in making final decisions. AI systems should not be completely relied on, regardless of any flawless track records or otherwise.

Technology Trust

As discussed in the third section of this report, the Turing Test was developed to assess the capabilities of AI. In 2018, the engineers at Google created a spoken-word NLP program called Duplex to integrate with its AI assistant. Its purpose is to make phone calls on a human's behalf, to talk to other humans, and answer questions in a natural way, all while sounding like a human (Leviathan, 2018). The program is able to search for the information needed as if a human were searching for it using, for example, Google. The AI assistant then calls, say, a restaurant to negotiate a time when the assistant's human will have a booked appointment. The program stutters, pauses, elongates certain vowels as if it had to think about what it is saying, and responds with alternative suggestions within its parameters after being given spoken information from a human hearing the commands.

For the purposes of this paper, at two different times, the authors asked an AI NLP program called 1558M about one of the research questions, in response to which the program had an interesting "opinion" of a negative and cautionary nature (Figure 36). This program was created to "play" with Open AI's new machine learning model (King, 2019). What is striking about these answers is that each is unique, meaning a search of the phrases finds no duplicates, but the phrasing and tone make it sound like they are from an informed source, with just enough minor evidence on the topic to seem believable. The program does not finish its last sentence, however, making it imperfect, but notable, nonetheless. Clearly, such AI capability holds much potential for assisting someone integrating 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 36. Two separate results from an AI called 1558M

***Source:* King (2019)**

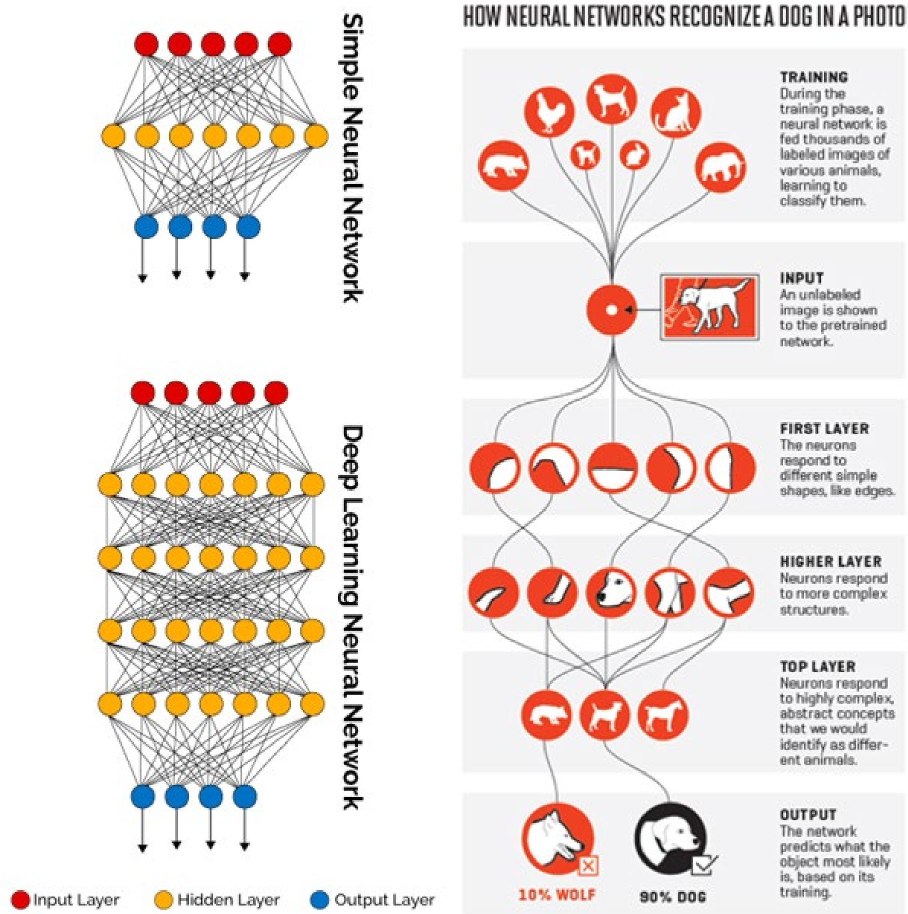
Explainable Reasoning

One of the limiting factors for AI adoption is being able to explain how the algorithm came up with its conclusion, which is critical for auditing (Knight, 2017). It would be negligent to use AI for military or financial purposes without the ability to trace how the decisions were made. Figure 37 shows how AI currently classifies information. For the multitude of training information that is used to create the program, the AI programs that turn out the desired result come up with their own way of navigating their layer complexities to create an output.

Fortunately for the DoD, the Defense Advanced Research Projects Agency (DARPA), already an organic element of the defense ecosystem, is leading the research into explainable AI (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)

Figure 35 visualizes the details of how a Deep Neural Network navigates its trained data to classify different pictures. This figure shows the images used to train an AI program on the left, and where in the neural network the associations of these trained data were used to classify an input to eventually reach a conclusion. Thus, if the DoD wanted to pursue human–machine partnerships in areas such as contracting, its organic system is providing the capability to do so.



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

Figure 37. Simple neural network compared to deep learning network

***Adapted from:* Golstein (2018); Parloff (2016)**

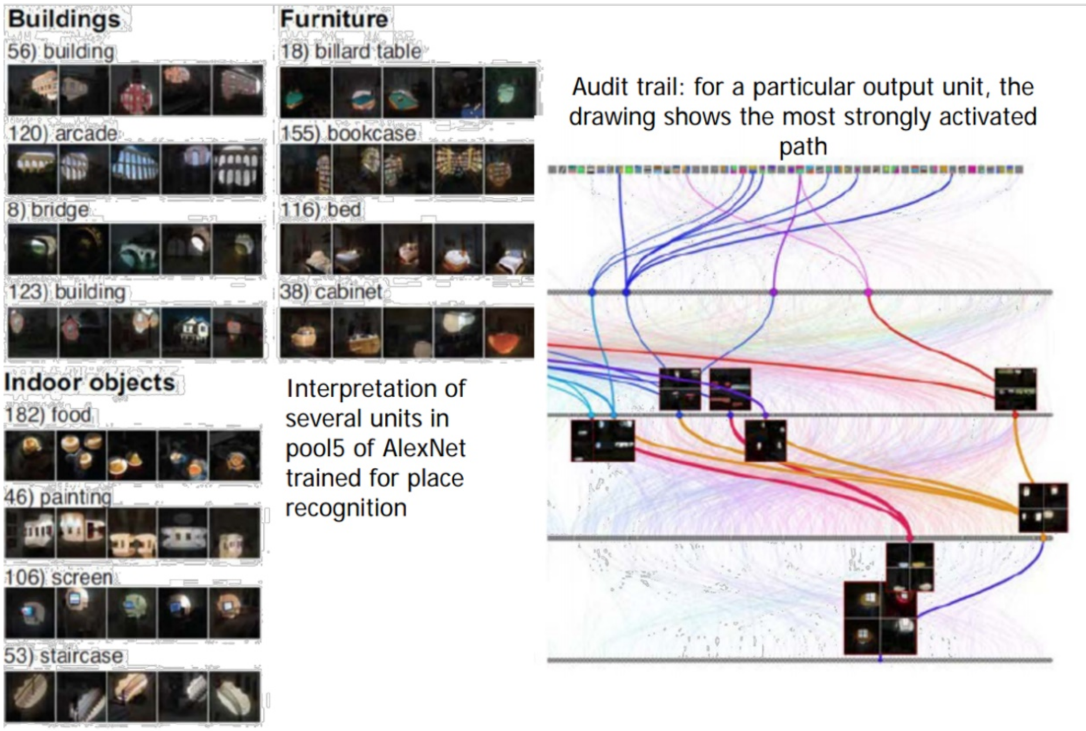


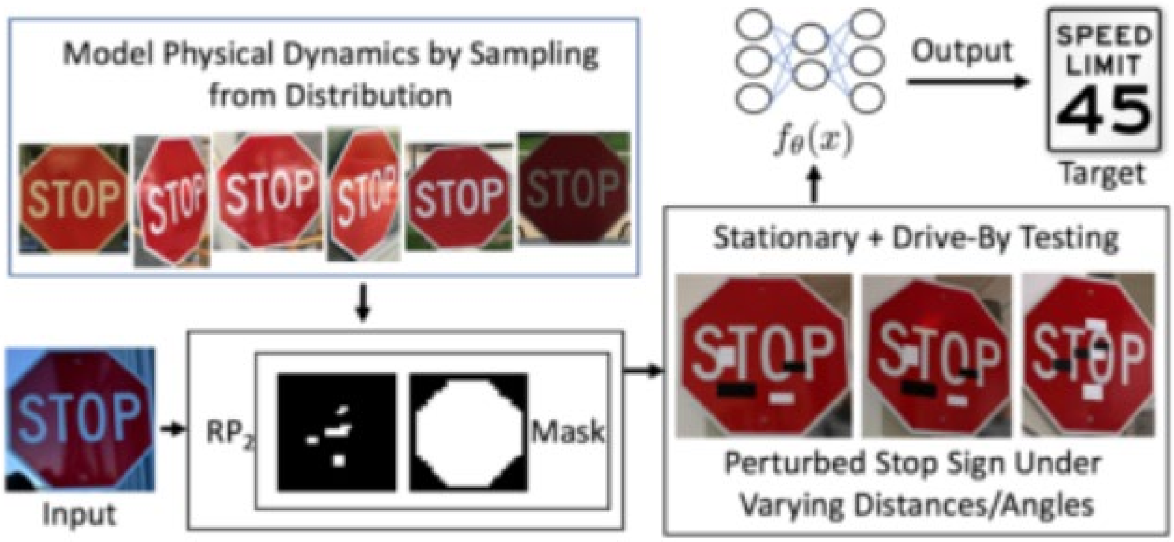
Figure 38. Visualization of explainable AI

Source: **DARPA (2019)**

Human–Machine Partnership

DARPA believes AI integration is critical as a human–machine symbiosis because sensor, information, and communication systems generate data at rates beyond which humans can assimilate, understand, and act on (DARPA, 2019). As was the case in the industrial revolution, machines are better at certain activities, and using machines for those activities frees humans to become productive in other areas. Humans and machines excel in separate areas of processing. Consider these comparisons between computers and humans: calculate vs. decide; compare vs. make judgments; apply logic vs. empathize; unaffected by tedious monotony vs. having preferences; deals with large data vs. intuitional focus on what is most important (Darken, 2019). And while AI performs well in some tasks, it works better with a human partner. Without proper controls, AI is a gullible learning system and can be vulnerable to being deceived by bad actors. Some studies show that AI can be fooled in a way that humans would not be due to human intuition. Other research has been able to fool a self-driving car into thinking a benignly tampered with stop sign was a speed limit sign (Figure 39), which would undoubtedly lead to collisions if the car were left unsupervised (Eykholt et al., 2018).

Many people are familiar with current intelligent machine partnerships that they, unknowingly, may experience on a daily basis. As discussed with its other applications, Google is the most popular search engine on the Internet because it provides better satisfaction to users than its competitors (Shaw, 2019). Google is so common as the preferred search engine that when someone talks about searching for something online, they refer to it as “Googling.” This is a good example of people interacting naturally with an AI system that uses Bidirectional Encoder Representations (BERT; Nayak, 2019). This is a technique that teaches a machine to not answer the user’s question based on individual words but, rather, on their meaning in the context of the question. For example, when asking what time it is right before lunch, the user is not asking for the actual time, but is really asking when they can eat; the outright answer would give the actual time, whereby the asker will deduce eating time, which was underlying meaning of the question. So-called self-driving vehicles provide another example of human interaction with intelligent machines. For the most part, the user sits in a supervisory role while the car takes over one of the most dangerous events in that person’s life and autonomously conducts all road tasks to drive (Darken, 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 39. AI system interpreting a stop sign

Source: **Eykholt et al. (2018)**

If contractors relied on an AI system to make all of the decisions for them, they would be susceptible to purposive misdirection by enemies introducing adversarial information for either competitive advantage or disruptions. Fraudsters can learn how to exploit computer algorithms, but only humans can judge the actions that result. What AI software can do, however, is easily extract data and clarify the content of contracts. It can quickly pull and organize the renewal dates and renegotiation terms from any number of contracts. It can let companies review contracts more rapidly, organize and locate large amounts of contract data more easily, decrease the potential for contract disputes and antagonistic contract negotiations, and increase the volume of contracts companies are able to negotiate and execute (Rich, 2018).

Case Study of Private Sector AI Application to Contracting

To compare solutions to the DoD acquisition problem, we examine similar circumstances in the U.S. private sector. Lawgeex is an example of a company that is applying the AI integration process in the private industry procurement world. It demonstrated that its AI software could outperform U.S. trained lawyers on an example contract aspect, the Non-Disclosure Agreement (NDA) with an average accuracy of 94% as compared to 85% for humans (Lawgeex, 2018). The study was conducted to respond to a common business problem in large companies that rely on contracts to engage with partners, suppliers, and vendors having an 83% dissatisfaction rate with their organization's contracting processes (Lawgeex, 2018). Another example is Icertis, a company that services large and commonly familiar companies such as 3M, Johnson & Johnson, and Microsoft to list a few (Icertis, n.d.-a). Icertis provides its customers with a cloud-based AI platform that learns from contracts provided by the client, along with control measures, to create and assist in contract setup; contract operations; governance, risk, and compliance; and reporting (Icertis, n.d.-a).

What makes this possible now, instead of when it was first theorized decades ago, is that industry is more accustomed to storing professional documents on a digitally accessible storage platform, whether local hard drives or the cloud (Betts & Jaep, 2017). Currently, the major hurdles that prevent a fully automated contract review and analysis process are nontechnical, such as the collection of contract performance data; publication of private contracts and their corresponding performance data; and changes in ethical constraints on computer usage in legal practice (Betts & Jaep, 2017). The authors of these obstacles also offer possible policy solutions to address them: start using contract management software that will be a forcing function to create data in an AI teachable format; expand copyright protection for vendors to protect their intellectual property; and create new rules to help mitigate AI risks to enable its ability to work (Betts & Jaep, 2017).

Cloud-Based AI

To understand how AI can be propagated throughout a system and update regulations and learn from multiple human teachers instantly, we look at the concept of cloud computing. The *speed of relevance* is a popular term in discussing DoD technology adoption. In the 2018 DoD Cloud Strategy, the term *cloud* refers to an offsite physical IT infrastructure. This external infrastructure communicates with a user's computer through the Internet to access data servers that store information and run operating systems, such as Microsoft Windows, which are centrally maintained. This means that every user always has the same software computing power and access to the most up-to-date software and is not limited by their organization’s IT professional talent or budget for new software. Organizations can get as much, or as little, access to what they need for projects and are unaffected by times of surging need or times of idleness, which currently add excess cost to DoD systems (Shanahan, 2018). The objective for DoD is to have AI-augmented rapid decision-making, in an environment where data is secure and visible for enhanced 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)

Acquisition Research Program
Graduate School of Defense Management
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net