RAND

Research Report

JAMES A. LEFTWICH, DAHLIA ANNE GOLDFELD, BRADLEY DEBLOIS,
CHAD HEITZENRATER, LUKE MUGGY, SHANNON PRIER, JOSHUA STEIER,
CHRISTOPHER E. MAERZLUFT, SYDNE J. NEWBERRY

Exploring Options to Improve Supply Chain Operations

A Review of Current Approaches and New Opportunities in Demand Forecasting, Robotic Process Automation, and Cyber Integrity

For more information on this publication, visit **www.rand.org/t/RRA1734-1**.

Published by the RAND Corporation, Santa Monica, Calif.

Cover: SSgt Renee Seruntine/U.S. Army, monsitj/Getty Images.

About This Report

The U.S. Air Force (USAF) has had long-standing concerns about its supply chains and their potential for degradation. Given the varied types of resources for which USAF manages the supply chains, the complexity of those supply chains, the reliance on the defense industrial base, and opportunities for degradation, the Headquarters U.S. Air Force Deputy Chief of Staff for Logistics, Engineering and Force Protection (HAF/A4) asked RAND Project AIR FORCE (PAF) to identify technology and investment options to mitigate supply chain degradation, focusing specifically on forecasting the demand for legacy aircraft spare parts, applying robotic process automation for supply chain management and execution, and mitigating risks associated with cyber integrity.

The RAND PAF research team employed different methodologies for each of the three lines of effort but developed use cases that cut across the different lines of effort to present findings and recommendations focused on cyber controls for integrity risks associated with the use of bots, as well as demand forecasting. We also highlighted opportunities for USAF to expand its current bot initiatives, including a use case that could have value in improving spare part demand forecasts for a unique class of parts and a method of approach to future investments, more generally, that could improve demand forecasting.

The research reported here was commissioned by HAF/A4 and conducted within the Resource Management Program of RAND Project AIR FORCE as part of a fiscal year 2022 project, "Investment Choices to Mitigate Supply Chain Degradation." This research should be of interest to USAF logisticians, as well as other Department of the Air Force (DAF) personnel responsible for cyber integrity risks and the current initiative to expand the application of robotic process automation across the DAF.

RAND Project AIR FORCE

RAND Project AIR FORCE (PAF), a division of the RAND Corporation, is the Department of the Air Force's (DAF's) federally funded research and development center for studies and analyses, supporting both the United States Air Force and the United States Space Force. PAF provides the DAF with independent analyses of policy alternatives affecting the development, employment, combat readiness, and support of current and future air, space, and cyber forces. Research is conducted in four programs: Strategy and Doctrine; Force Modernization and Employment; Resource Management; and Workforce, Development, and Health. The research reported here was prepared under contract FA7014-22-D-0001.

Additional information about PAF is available on our website: www.rand.org/paf/

This report documents work originally shared with the DAF on September 30, 2022. The draft report, dated September 2022, was reviewed by formal peer reviewers and DAF subject-matter experts.

Acknowledgments

We thank Major General Linda Hurry for sponsoring this work and providing support throughout the year. Similarly, her staff provided access to information and insights as we conducted the research. Specifically, we thank Kim Brown, Colonel Brad Coley, Colonel Jim Hartle, Steve Martinez, Lieutenant Colonel Travis Bohanan, Major Michael Ingram, Major Michael J. Hester, William W. Wenzel, and Captain Nicholas L. Spivey. Additionally, several members of the HAF/Resource Integration Directorate (A4P) staff, specifically Carlo de Guzman and Abigail Strobell, provided insights and updates on the HAF/A4's cybersecurity efforts.

Special thanks go to Steve Gray, director of the 448th Supply Chain Management Wing, for his early engagement and providing us access to other members of the 448th Supply Chain Management Wing who provided us insights and information. Members of the Enterprise Supply Chain Analysis, Planning, and Execution (ESCAPE) program office and other members in the 420th Supply Chain Management Squadron met with us regularly to provide information and data and answered questions, as did Rich Moore, Air Force Sustainment Center Lead Analyst. At the Air Force Life Cycle Management Center, Joseph Besselman, Chief Disruption Officer for the Command, Control, Communications, Intelligence and Networks Directorate and Business Enterprise Systems Directorate, and Matthew Roberts, from the Robotic Process Automation Center of Excellence, were very gracious with their time to help us better understand initiatives associated with robotic process automation.

At RAND, we thank John Drew, Tim Conley, and Dan Romano for their contributions to the research. Also at RAND, thanks go to Anna Jean Wirth, Tom Light, Li Ang Zhang, and Yusuf Ashpari for meeting regularly to share updates and insights on related efforts they were leading. Finally, we are grateful to Elvira Loredo and Caolionn O'Connell, who provided thorough and thoughtful reviews that improved the quality of the report.

That we received help and insights from those acknowledged above should not be taken to imply that they concur with the views expressed in this report. We alone are responsible for the content, including any errors or oversights.

Summary

Issue and Approach

The U.S. Air Force (USAF) has had long-standing concerns about its supply chains and their potential for degradation. Given the varied types of resources for which USAF manages the supply chains and opportunities for degradation, the RAND Project AIR FORCE research team identified technology and investment options to mitigate supply chain degradation, focusing specifically on forecasting the demand for legacy aircraft spare parts, applying robotic process automation (RPA),[1] and mitigating risks associated with cyber integrity. We conducted extensive literature reviews, held discussions with subject-matter experts, and employed various analytical methods. For demand forecasting, these methods included analysis of recent forecast accuracy data to identify drivers of forecast error and an assessment of ongoing efforts to address known issues. For RPA, our analyses included a characterization of bot development and identification of potential application areas in the logistics, engineering, and force protection (A4) community. For cyber integrity, the analysis involved applying a mission assurance approach to identify potential risks and mitigations.

Key Findings and Conclusions

Demand Forecasting

- Primary drivers of demand forecast errors, such as propulsion systems and low demand for expensive parts, are well known to personnel who study the problem.
- Decades of research on demand forecasting suggest there are a variety of methods to forecast spare parts demand, although there is not a one-size-fits-all best approach.
- USAF's migration to a commercial enterprise resource planning system for demand forecasting is already showing promise.[2]
- It is unclear whether demand forecast accuracy is resulting in aircraft downtime.

Application of Robotic Process Automation to Improve Supply Chain

- The USAF A4 community's current approach to bot implementation does not allow it to fully leverage the potential of bots.

---

[1] RPA is a term used to describe the employment of a software program or application, referred to as a *bot*, to automate a process that is manual, repetitive, and typically manpower intensive.

[2] Through the Enterprise Supply Chain Analysis, Planning, and Execution (ESCAPE) program office, USAF is implementing PTC's service parts management software solution, identified by Accenture as a best-in-class solution for large, complex supply chains and in use today by such companies as Airbus, Boeing, and Quantas.

- Questions remain about whether USAF personnel possess the technical expertise to fully leverage bot technology, and the data suggest this concern is warranted.
- Unified direction and guidance could help USAF maximize the potential of bots.

Cyber Tampering

- Potential vectors for integrity attacks within the supply chain include the software supply chain, software vulnerabilities, and credential-based attacks.
- Risks related to both development and operation of bots underscore the need to consider cyber risk across the bot life cycle and incorporate training and best practices.
- In addition to existing processes, risk-based analysis can drive cybersecurity, engineering, and mission-execution decisionmaking for technology under consideration.
- Current Headquarters U.S. Air Force Deputy Chief of Staff for Logistics, Engineering and Force Protection (HAF/A4) risk management framework (RMF) controls are primarily focused on detection and are unlikely to sufficiently mitigate focused, tampering-based mischief.

Recommendations

Demand Forecasting

- USAF should maximize the potential benefits offered by the investment already made in ESCAPE. However, USAF should analyze the value of demand forecast improvements to supply chain performance prior to making additional investments.
- If additional investment in forecast accuracy improvement is warranted, the 448th Supply Chain Management Wing should target specific areas of improvement, such as expanded causal analysis for parts with intermittent, infrequent, and highly variable demand.

Application of Robotic Process Automation to Improve Supply Chain

- USAF should expand the application of bots within the A4 community, including processes and data integration not currently accomplished. We provide a suggestion with our bot example.
- USAF should work with the USAF lead for RPA to establish standards for centralized development and management of bots and should advocate for funding for increased security measures.

Cyber Tampering

- USAF should consider mitigation approaches for integrity attacks identified by this analysis, especially in bot implementation, as a complement to existing cybersecurity controls.

- USAF should continue to evaluate cyber risks in context by implementing a process for considering how threats, vulnerabilities, and consequences to missions change as new systems, technologies, and information-handling methods are considered and implemented.
- USAF should employ best practices for executing risk-based processes (e.g., Operationally Critical Threat, Asset, and Vulnerability Evaluation [OCTAVE] Allegro),[3] such as engaging subject-matter experts on the value of information assets to the mission and complementing HAF/A4 RMF with a cross-functional approach.

[3] OCTAVE Allegro is a risk-based cybersecurity assessment methodology that provides a means to examine the threats, vulnerabilities, and consequences of attack within a system. As used in this research, OCTAVE provided a structured approach to evaluating the risk to mission posed by cybersecurity threats to information and evaluating alternative mitigations.

Contents

Figures and Tables

Figures

Tables

Chapter 1

Introduction

In January 2020, the World Health Organization (WHO) revealed that unexplained illnesses in Wuhan, China, were the result of coronavirus disease 2019 (COVID-19). By March 11, 2020, given the rising number of infections and the rate of spread, the WHO declared COVID-19 a global pandemic.[4] Less than two weeks later, analysts were speculating on how COVID-19 could affect U.S. Department of Defense (DoD) supply chains; concerns focused on the impacts of production shutdowns and labor shortages, as well as single sources of production and supply of unique and critical items folding under the weight of financial distress leading to bankruptcy.[5] Although concerns about the fragility of DoD supply chains were not new, COVID-19 introduced yet another unexpected disruption that turned out to have major consequences, particularly for key machine subcomponents like chips and microprocessors (which already faced problems) but also for items like personal protective equipment and unexpected commodities like toilet paper.

DoD had already expressed concern about supply chain resiliency tied to operating in a conflict against a near-peer adversary capable of creating a contested battlespace that could disrupt lines of communication and supply and support to combat forces.[6] This concern has been recognized by the U.S. government. For example, in 2017, former President Donald Trump issued an executive order requiring the Secretary of Defense to conduct a holistic review of the threats to the DoD supply chain and industrial base.[7] On February 14, 2021, President Joe Biden signed Executive Order 14017, which led to a study focused on the supply chains that are critical to national security and the everyday functioning of the U.S. population. Similar to the 2017 study, the research objectives were to identify vulnerabilities and secure four key supply chains, while setting the groundwork to instantiate mitigation strategies for the long term.[8]

As a service, the U.S. Air Force (USAF) has also long been focused on this problem, which has become exponentially more complicated in recent years and decades. There are many reasons for this, including the fact that USAF weapon platforms are increasingly integrated, USAF faces a variety of obsolescence problems, and USAF struggles to illuminate its own supply chains beyond the first- and

[4] Centers for Disease Control and Prevention, "CDC Museum COVID-19 Timeline," webpage, last reviewed March 15, 2023.

[5] Aaron Mehta, "How Coronavirus Could Impact the Defense Supply Chain," Defense News, March 20, 2020.

[6] Frank Wolfe, "Joint Warfighting Concept Assumes 'Contested Logistics,'" *Defense Daily*, October 6, 2020.

[7] Interagency Task Force in Fulfillment of Executive Order 13806, *Assessing and Strengthening the Manufacturing and Defense Industrial Base and Supply Chain Resiliency of the United States*, September 2018.

[8] The White House, *Building Resilient Supply Chains, Revitalizing American Manufacturing, and Fostering Broad-Based Growth: 100-Day Reviews Under Executive Order 14017*, June 2021.

second-tier suppliers.[9] RAND Project AIR FORCE (PAF) has conducted a long series of studies, dating back to the 1960s, on the challenges associated with ensuring secure and stable USAF supply chains.

In addition to the U.S. government renewing its focus on supply chain security, the commercial sector has also been independently attacking this problem. The pandemic has simply magnified already long-standing and growing issues stemming from such factors as climate change and increasing global interdependence.

For the purpose of this report, we used the Merriam-Webster definition of *supply chain*: "the chain of processes, businesses, etc. by which a commodity is produced and distributed: the companies, materials, and systems involved in manufacturing and delivering goods."[10] For USAF, this chain encompasses multiple commodities (referred to as *classes of supply*), which reach end users through different mechanisms. First, the industrial base produces and delivers commodities directly to USAF. Second, other DoD agencies acquire, manage, and ultimately deliver commodities to USAF. An important example of this is jet fuel, the supply chain of which is mostly handled by the Defense Logistics Agency (DLA). Finally, organic USAF organizations manage and deliver commodities to other USAF consumers. For example, the 448th Supply Chain Management Wing (SCMW) and the 635th Supply Chain Operations Wing within the Air Force Sustainment Center (AFSC) are responsible for such items as spare parts, equipment, repair capacity, and fuels management within USAF. For the purposes of this report, we defined *supply chain degradation* as any activity that disrupts the timely production and distribution of a commodity or compromises its quality.

A preliminary literature search on analysis of USAF and DoD supply chains revealed multiple studies by RAND alone that spanned a wide variety of important supply chain topics. Some research focused on the supply chain managed by USAF or DoD organizations.[11] Other research focused on the industrial base side of the equation.[12] Some research focused on enterprise-level challenges,[13] and other efforts focused more deeply on single commodities.[14]

---

[9] Supply chain illumination is a supply chain risk management process in which an organization will conduct a thorough review of its suppliers, or supply chains for particular parts, to identify vulnerabilities. USAF has launched an illumination effort to improve its awareness of risks to its supply chains. See Michele Donaldson, "Team Illuminates Supply Risks That Impact Defense," Eglin Air Force Base, August 31, 2022.

[10] "Supply Chain," webpage, Merriam-Webster, undated.

[11] See, for example, Eric Peltz, Amy G. Cox, Edward W. Chan, George E. Hart, Daniel Sommerhauser, Caitlin Hawkins, and Kathryn Connor, *Improving DLA Supply Chain Agility: Lead Times, Order Quantities, and Information Flow*, RAND Corporation, RR-822-OSD, 2015; and G. J. Feeney and C. C. Sherbrooke, *Systems Analysis and Supply Management*, RAND Corporation, RM-4054-PR, 1964.

[12] Anthony G. Bower and Steve Garber, *Statistical Forecasting of Bankruptcy of Defense Contractors: Problems and Prospects*, RAND Corporation, MR-410-AF, 1994.

[13] Nancy Y. Moore, Elvira N. Loredo, Amy G. Cox, and Clifford A. Grammich, *Identifying and Managing Acquisition and Sustainment Supply Chain Risks*, RAND Corporation, RR-549-AF, 2015; and Caolionn O'Connell, Elizabeth Hastings Roer, Rick Eden, Spencer Pfeifer, Yuliya Shokh, Lauren A. Mayer, Jake McKeon, Jared Mondschein, Phillip Carter, Victoria A. Greenfield, and Mark Ashby, *Managing Risk in Global Supply Chains*, RAND Corporation, RR-A425-1, 2021.

[14] See, for example, Caolionn O'Connell, Bryan Boling, Jonathan Balk, James R. Broyles, and Monika Cooper, *Hidden Disruptions to the Supply Chain: Resistance Is Futile*, RAND Corporation, 2021, Not available to the general public; Dwayne M. Butler, Anthony Atler, Stephen M. Worman, Lily Geyer, and Bonnie Magnuson, *Identifying Efficiencies in the Supply Chain for*

Scope of This Effort

Clearly, supply chain degradation can come from many sources, and, thus, securing and increasing the reliability of the myriad supply chains that ultimately deliver commodities to end users is vastly complex. Mitigation approaches correspondingly vary. Generally speaking, DoD, which relies on the industrial base, worries about both security and reliability of products, and, to this end, it has to deal with diminished sources of supply for their production lines or service (which includes both raw materials and subcomponents), labor and workforce disruptions, counterfeits, malicious tampering of electronics goods, bankruptcy, government regulation, and more. For those parts of supply chains over which USAF and DoD have more control, stakeholders seek solutions to mitigate disruptions resulting from poor inventory management and planning, cyber intrusion, poor repair planning, inadequate tracking of external supply chains, unstable budgets, and others. There are a variety of general approaches that USAF can take to attack these challenges, including economic incentives, improved internal processes and workforce management, and new technologies. To this end, Headquarters U.S. Air Force Deputy Chief of Staff for Logistics, Engineering and Force Protection (HAF/A4) asked RAND PAF to identify technology and investment options to mitigate supply chain degradation.

Technology, of course, still represents a very broad set of options that could potentially mitigate USAF-specific supply chain concerns. Thus, HAF/A4 asked us to focus on near-term options related to the parts of the supply chain that USAF directly controls. In particular, HAF/A4 identified three main lines of effort, which subsume the remainder of this report. They are the following:

- **Demand forecasting**: How does USAF forecast demand today and how could demand forecasting be improved?
- **Bots**: How are bots used in USAF today and how can they be used to improve supply chain effectiveness and efficiency?
- **Cyber integrity**: How can USAF identify and mitigate integrity-based risk? Because HAF/A4 is very concerned about detecting incidents of data tampering, we looked for approaches USAF can take to detect and prevent this problem.[15]

Each line of effort involved different methodologies, which are described in their respective chapters. During our research, interdependencies and links between the three separate research efforts emerged. For example, we came to understand that a fundamental challenge associated with forecasting the demand for spare parts is that the data needed to potentially improve forecasting are not easy to compile. This finding sets the stage for the possible use of a bot that can assemble the necessary dataset and even do analysis, which might further inform demand forecasts. Not only does this help to reveal a major bot-related finding—that HAF/A4 should expand its thinking about bots to cross-functional realms that might even harness machine intelligence—but it also inspired us to

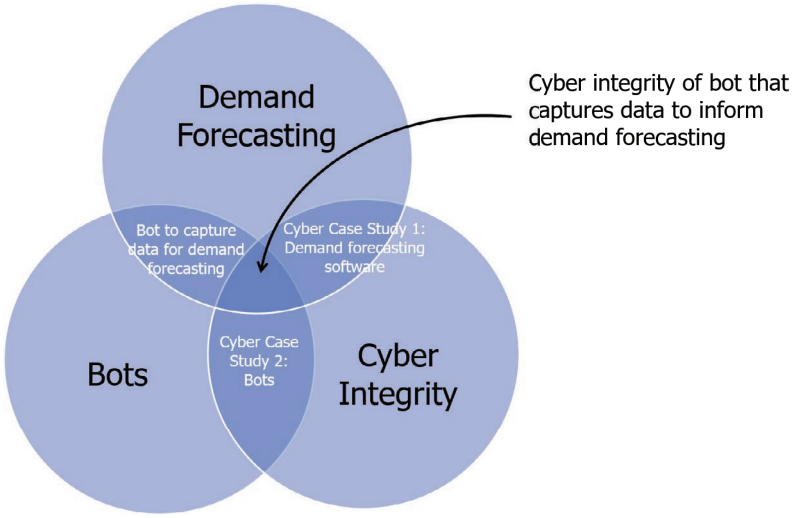
---

Training Ammunition: Methods, Models, and Recommendations, RAND Corporation, RR-952-A, 2016; and Elvira N. Loredo, John F. Raffensperger, and Nancy Y. Moore, *Measuring and Managing Army Supply Chain Risks: A Quantitative Approach by Item Number and Commercial Entity Code*, RAND Corporation, RR-902-A, 2015.

[15] The sponsor expressed *cyber-tampering events* as those that would be classified as mischief rather than mayhem. Examples included such things as manipulating data or algorithms that would generally go unnoticed during daily operations.

suggest a detailed bot use case to HAF/A4. Our research also revealed the need for caution in implementing bots, because they introduce a major, but easy to overlook, potential risk of cyber vulnerabilities. For this reason, we used bots as a springboard to illustrate a cybersecurity approach to minimize access points for cyber tampering. Figure 1.1 visualizes some of the connections between the three lines of effort just described.

Figure 1.1. Relationship Between Three Lines of Effort Addressed in This Report



Along with the scoping guidance described previously, the sponsor asked that we be cognizant of and consider how our recommendations would tie to ongoing efforts within USAF that touch on our three lines of effort.[16] Thus, our general approach for this project included an effort to connect with a broad community of stakeholders that might have interest and equities in the outcomes of our research.

Organization of This Report

The remainder of this report is organized as follows:

- Chapter 2 presents our assessment of how USAF could improve the accuracy of legacy aircraft spare parts demand forecasting.
- Chapter 3 provides the results of our bots analysis and includes a suggested bot use case that connects to demand forecasting.
- Chapter 4 presents a risk-based cybersecurity approach based on the Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) Allegro process and our analysis of mitigation options to address cyber-tampering vulnerabilities. We illustrate this approach using three different use cases. The first focuses on cyber tampering of demand forecasting

[16] For example, for demand forecasting, USAF had just migrated from a legacy system to a new commercial product. Similarly, USAF and HAF/A4 have also started an initiative to gather ideas on how bots could be used to eliminate repetitive labor-intensive tasks being performed by service members.

tools and systems. The second employs the OCTAVE Allegro process to examine ways to harden future bots against cyber tampering. The third is intended to suggest specific ways to improve the cybersecurity of the bot proposed in Chapter 3.

- Chapter 5 summarizes the most-important observations and findings of our research and presents our recommendations.
- Appendix A is an annotated bibliography of research related to our demand forecasting analysis.
- Appendix B includes additional details associated with our bots research.
- Appendix C provides additional details associated with our cyber-tampering research.
- Appendix D provides the rationale and motivation behind the selection of OCTAVE Allegro for our cyber analysis.
- Appendix E presents OCTAVE risk and mitigation analysis details specific to our three use cases.

Chapter 2

Evolution of Spare Part Demand Forecasting in USAF

The Air Force Materiel Command (AFMC) manages materiel support for USAF, and the command's AFSC is responsible for managing secondary inventory, including spare parts. Specifically, AFSC's 448th SCMW is responsible for the planning and execution of spare part requirements for a wide variety of systems, including aircraft; engines; intercontinental ballistic missiles; and space command, control, communication, and intelligence (C3I) equipment.[17] The 448th SCMW and its 3,000 civilian and military personnel conduct spare parts demand planning, supply planning, and inventory optimization to ensure that the complex USAF supply system delivers the right part to the right place at the right time.[18] As operating and support costs have increased over the years, USAF has continually looked for ways to improve the efficiency of its resource use.[19] USAF spends billions of dollars annually on spare parts, so potential improvements to inventory management continue to receive considerable attention. *Demand forecasting*, a principal component of inventory management, generally refers to the act of predicting future customer demands so inventory managers can develop inventory requirements to satisfy demands when they occur. In the USAF supply chain enterprise, a *demand* is defined as an indication of a requirement, a requisition, or a similar request for an item of supply or individual item to support the repair and maintenance of weapon systems, reparable end items, and equipment.[20] Inaccurate forecasting of these demands can lead to either excess inventory with associated cost or shortfalls that could affect mission effectiveness. As a result, demand forecasting accuracy is under continuous review by USAF,[21] and it is reported semiannually to the Office of the Secretary of Defense (OSD).[22]

[17] "448th Supply Chain Management Wing," webpage, Tinker Air Force Base, undated.

[18] "448th Supply Chain Management Wing," undated.

[19] Thomas Light, Michael Boito, Tim Conley, Larry Klapper, and John Wallace, *Understanding Changes in U.S. Air Force Aircraft Depot-Level Reparable Costs over Time*, RAND Corporation, 2018, Not available to the general public.

[20] AFMC Manual 23-101 Volume 1, *Materiel Management General D200A/N Information*, Department of the Air Force, November 17, 2016.

[21] 420th Supply Chain Management Squadron (SCMS) personnel, discussion with authors, January 3, 2022.

[22] DoD, *Supply Chain Metrics Guide*, 3rd ed., 2021.

U.S. Department of Defense Emphasis on Improving Demand Forecasting

There has been considerable effort across DoD to improve supply chain management. From 2007 to 2009, the Government Accountability Office (GAO) conducted audits of the military services' and DLA's spare part inventories and found significant excess of some items across all services and substantial shortfalls for certain items. GAO recommended that the services improve demand forecasting, reduce inventory, evaluate parts retention decisions, and establish metrics for tracking cost efficiency of inventory management.[23] Around the same time, the fiscal year (FY) 2010 National Defense Authorization Act (NDAA) included a section directing a series of plans requiring DoD to rightsize its secondary items inventory and continue effective and efficient materiel support to warfighters.[24] The NDAA laid out eight required elements for DoD to review to rightsize the inventory: demand forecasting procedures, total asset visibility, excess on-order secondary inventory, economic retention requirements, contingency retention requirements, a potential shift to direct vendor delivery for some items, no recurring demand items, and additional disposal.[25] In response, DoD released the Comprehensive Inventory Management Improvement Plan (CIMIP) in November 2010.[26] The CIMIP included a subplan focused on each of these eight elements.

The subplan on demand forecasting expanded on an existing 2009 OSD initiative to improve demand forecasting throughout the life cycle of secondary items managed by the services and DLA. The goal of the demand forecasting subplan was to improve the prediction of future demand so that inventory requirements more accurately reflect actual needs. To accomplish this goal, the subplan proposed five actions: the identification of improved methods for demand forecasting, implementation of standard metrics to assess forecast accuracy and bias, expansion of collaborative forecasting, improvement in the determination of inventory levels for low-demand items, and reduction of the investment risk of consumable items initially entering the inventory.[27]

In the following years, OSD commissioned three studies on demand forecasting to satisfy specific parts of the CIMIP, and the services and DLA all developed specific improvement plans. Following subsequent reviews in 2015 and 2017, the GAO removed DoD inventory management, which included demand forecasting, from its list of high-risk areas for the first time since 1990, citing sufficient improvement across all five high-risk criteria: leadership commitment, capacity, action plan,

---

[23] GAO, *Defense Inventory: Defense Logistics Agency Needs to Expand on Efforts to More Effectively Manage Spare Parts*, GAO-10-469, May 11, 2010; GAO, *Defense Inventory: Army Needs to Evaluate Impact of Recent Actions to Improve Demand Forecasts for Spare Parts*, GAO-09-199, January 12, 2009; GAO, *Defense Inventory: Management Actions Needed to Improve the Cost Efficiency of Navy's Spare Parts Inventory*, GAO-09-103, December 12, 2008; and GAO, *Defense Inventory: Opportunities Exist to Save Billons by Reducing Air Force's Unneeded Spare Parts Inventory*, GAO-07-232, April 27, 2007.

[24] Public Law 111-84, National Defense Authorization Act for Fiscal Year 2010, October 28, 2009. *Secondary* refers to items of supply, including reparable components, subsystems and assemblies, consumable repair parts, bulk items and materials, subsistence, and expendable end items (i.e., clothing and other personal gear), that are not defined as principal items. A principal item would be a weapon system (Department of Defense Manual 4140.01 Volume 2, *DoD Supply Chain Materiel Management Procedures: Demand and Supply Planning*, U.S. Department of Defense, 2018).

[25] A *no recurring demand item* is defined as an item for which demand is low or sporadic.

[26] GAO, *DOD's 2010 Comprehensive Inventory Management Improvement Plan Addressed Statutory Requirements, but Faces Implementation Challenges*, GAO-11-240R, January 7, 2011.

[27] GAO, 2011.

monitoring, and demonstrated progress.[28] The GAO credited USAF with institutionalizing the demand forecast accuracy (DFA) metric and efforts to review 200 items per quarter.

Demand Forecasting in USAF

With the context of the importance of demand forecasting at the DoD level, we turn to the USAF-specific demand forecasting approaches and metrics. Long before the recent emphasis on improving forecast accuracy, USAF depended on a reliable supply of spare parts and, thus, has a rich history of approaches to forecasting demand. A persistent theme since the earliest studies of demand for spare parts is that demand is inherently uncertain.[29] The sources of the uncertainty are generally grouped into two categories: (1) inherent statistical uncertainty arising from the stochastic nature of peacetime operations and (2) external fluctuations induced by contingency operations.[30]

Over time, two approaches have emerged to mitigate the uncertainty. The first approach is to improve demand forecasting via improved models, analysis of drivers, and increased data collection and aggregation. The second approach is to focus on strategies and processes that are adaptive and robust against inevitable uncertainty. The emphasis placed on each approach has fluctuated over time depending on world events and fiscal realities. The general consensus is to predict as accurately as possible, while retaining sufficient resilience to accommodate uncertainty.[31] With the recent emphasis on prediction at the DoD level and decreases in the cost of data storage and emerging methods in data science (in particular, the application of machine learning [ML] and artificial intelligence [AI]), there has been renewed interest in exploring more sophisticated techniques for demand forecasting. Before exploring those options, we describe the legacy demand forecasting methodology used by USAF, present our analysis of recent demand forecasting performance, and describe a major change to that methodology that is already in motion.

Legacy Demand Forecasting

The 448th SCMW uses the Requirements Management System (RMS) to track needs and project requirements to enable maintenance and repair of USAF weapon systems. The Secondary Item Requirements System, a legacy information technology (IT) system aligned under the RMS and referred to as *D200A*, has been the primary planning component of RMS used to compute spare part requirements on an aggregate basis and apply worldwide assets to those requirements.[32] One key function of D200A is forecasting the demand for spare parts. The total demand for a part is driven

[28] GAO, *High Risk Series*, GAO-15-290, February 11, 2015; and GAO, *Progress on Many High-Risk Areas, While Substantial Efforts Needed on Others*, GAO-17-317, February 15, 2017.

[29] For the earliest RAND report on the topic, see B. B. Brown and M. A. Geisler, *Analysis of the Demand Patterns for B-47 Airframe Parts at Air Base Level*, RAND Corporation, RM-1297, 1954. See also Bernice B. Brown, *Characteristics of Demand for Aircraft Spare Parts*, RAND Corporation, R-292, 1956.

[30] James S. Hodges and Raymond A. Pyles, *Onward Through the Fog: Uncertainty and Management Adaptation in Systems Analysis and Design*, RAND Corporation, R-3760-AF/A/OSD, 1990.

[31] Hodges and Pyles, 1990.

[32] AFMC Manual 23-101 Volume 1, 2016.

primarily by base-level maintenance (referred to as *organization and intermediate maintenance* [OIM]) and depot-level maintenance (DLM).

OIM requirements are driven by both scheduled and unscheduled maintenance. For base-level demand, the basic model in D200A calculates straight-line forecast factors based on past use.[33] It calculates the demand rate as the total number of demands in the past eight quarters divided by the total flying hours in the past eight quarters. The future requirement is then predicted by multiplying the projected flying hours in some future period by the demand rate. This is referred to as the *eight-quarter moving average forecast*. D200A also includes alternative model options, including a four-quarter moving average, exponential smoothing, and a regression technique. In addition, the number of flying hours is only one of several options for program data that can drive the need for spare parts. Other types of OIM program data include inventory, sorties, drone recoveries, and ammunition.[34] However, the current system allows the use of only one type of program data at a time, and flying hours is the most commonly used.[35] The model assumes a perfectly linear relationship between flying hours and demands, and past research has highlighted the possibility that such a relationship does not always hold.[36]

DLM requirements are driven by three program types: programmed depot maintenance (PDM), engine overhauls (EOHs), and next higher assembly management of items subject to repair (NHA MISTR). PDM involves a schedule-based inspection and correction of defects that requires skills, equipment, or facilities not normally possessed at operating locations. EOHs include the disassembly and inspection, repair, replacement, and servicing of engines. NHA MISTR requirements are driven by the DLM of aircraft-level parts.

The 448th SCMW assesses DFA to identify problems, track progress toward goals, and report performance to USAF leaders and DoD. Using guidance from DoD as part of the CIMIP and now outlined in the DoD *Supply Chain Metrics Guide*,[37] USAF uses two metrics to assess forecast accuracy: DFA and bias.[38] DFA is a measure of the percentage difference between the forecasted and actual demand. It is calculated as 1 minus the absolute difference between actual demand and forecasted demand, weighted by the dollar value of each part, divided by the dollar weighted value of the actual demand. USAF uses latest acquisition cost to define the dollar value for each part.[39]

---

[33] AFMC Manual 23-101 Volume 5, *Equipment Specialist Data and Reports (D200A, D200N)*, Department of the Air Force, December 15, 2021.

[34] AFMC Manual 23-101 Volume 5, 2021.

[35] Joshua D. DeFrank, "A Condition Based Maintenance Approach to Forecasting B-1 Aircraft Parts," Air Force Institute of Technology, March 3, 2017.

[36] Craig C. Sherbrooke, *Using Sorties vs. Flying Hours to Predict Aircraft Spares Demand*, Logistics Management Institute, 1997; and Thomas R. O'Neal, *Sortie-Based Aircraft Component Demand Rate to Predict Requirements*, Air Force Institute of Technology, March 2020.

[37] DoD, 2021.

[38] There are a variety of forecast accuracy metrics commonly used in literature and practice. For an overview, see Rob J. Hyndman, "Another Look at Forecast-Accuracy Metrics for Intermittent Demand," *Foresight*, No. 4, June 2006. DFA, as defined here, is a type of percentage error. Percentage errors have the advantage of being scale independent and, thus, useful for comparing forecast performance. They can be problematic for intermittent demand because a zero-demand case results in an undefined value. Because DFA typically aggregates across parts, this drawback is somewhat alleviated for the purpose of this analysis.

[39] Latest acquisition cost is the most recent acquisition cost but could be from several years in the past.

The equation for DFA is shown here:[40]

$$DFA = \left[1 - \frac{\sum_{NIIN}(|predicted\ demand - actual\ demand| \times dollar\ value)}{\sum_{NIIN}(actual\ demand \times dollar\ value)}\right] \times 100.$$

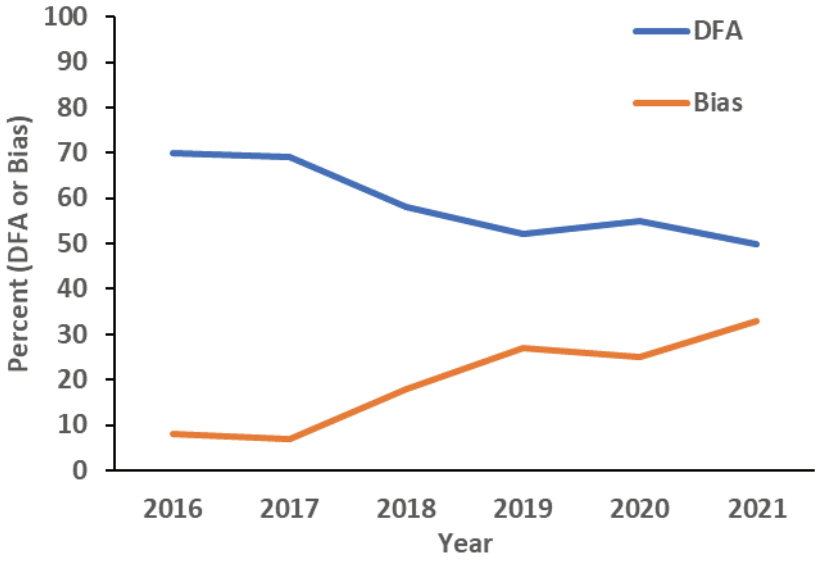
The other metric, bias, is a measure of the over or under percentage difference between the forecasted and actual demand. The equation is shown here:

$$Bias = \frac{\sum_{NIIN}([predicted\ demand - actual\ demand] \times dollar\ value)}{\sum_{NIIN}(actual\ demand \times dollar\ value)} \times 100.$$

DFA provides information on the overall magnitude of the forecasting error, and bias provides insight into whether that error tends to be a result of over- or under-forecasting. Over-forecasting results in a positive bias because predicted demand is much greater than actual demand. Under-forecasting results in a negative bias because predicted demand is much lower than actual demand.

Figure 2.1 shows the DFA and bias for all USAF-managed spare parts from 2016 to 2021 (as calculated in March of each year). The 448th SCMW sets a target DFA of 71 percent. In 2016–2017, they nearly met that target, but from 2017 to 2021, DFA decreased from 70 percent to 50 percent. Bias has increased from 10 percent to 30 percent over the same period, indicating that the reduction in DFA is primarily a result of over-forecasting. To improve forecast accuracy, it is important to understand the sources of forecast error so that the investment of time and resources can be applied to the areas of largest potential benefit. In the next section, we examine the drivers of recent forecast error.

Figure 2.1. Demand Forecast Accuracy and Bias, 2016–2021



[40] The national item identification number (NIIN) is a nine-digit number used to identify a specific part. Each part in the supply system will have a unique NIIN (DLA, undated).

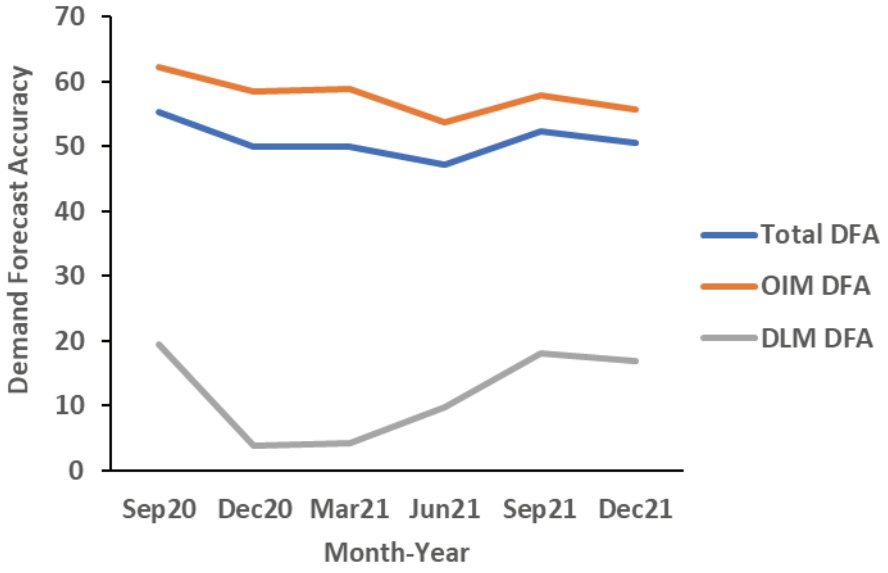
SOURCE: Produced using data from USAF, 420th Supply Chain Management Squadron, *AF DFA and Bias with IMR Charts,* PowerPoint presentation, July 11, 2022.
NOTE: These values were derived from a Microsoft PowerPoint briefing provided by 420th SCMS personnel (USAF, 420th Supply Chain Management Squadron, 2022).

Drivers of Recent Demand Forecast Accuracy and Bias

To assess drivers of recent changes in DFA and bias, we obtained more-detailed data from the 448th SCMW for 2020–2021. The data provided included quarterly reports from September 2020 to December 2021, generated by the Forecast Analysis Comparison Tool (FACT) Plus, the web application that 448th SCMW personnel use to track DFA. Each report included detailed information about each USAF-managed part, including group, squadron, latest acquisition cost, total annual forecast (as forecasted the prior year), and total annual demand. It also included annual forecast and demand for OIM and DLM. This granular level of detail enabled a decomposition of DFA and bias in a variety of ways, including by OIM versus DLM, part type, and demand quantity, as discussed in the following sections.

Figure 2.2 shows the total, OIM, and DLM DFA from September 2020 to December 2021. Each quarter represents an annual DFA, so the figure essentially shows a rolling annual calculation. The total DFA is approximately 50 percent during this period, the same as in Figure 2.1. However, when analyzing OIM and DLM separately, we can see that OIM DFA is higher (at approximately 60 percent), while DLM DFA is below 20 percent. OIM accounts for more than 75 percent of the dollar value of total demand, so it has a larger impact on total DFA. OIM DFA is still below the 71 percent target, while DLM is significantly below that target.

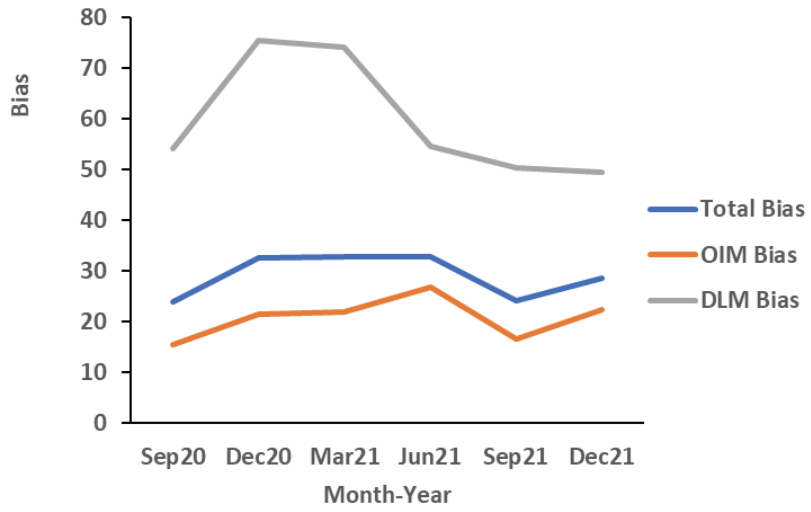
Figure 2.2. Demand Forecast Accuracy—Total, Organization and Immediate Maintenance, and Depot-Level Maintenance



SOURCE: Produced using data from USAF, Forecast Analysis Comparison Tool, database, 2020–2021.

Figure 2.3 shows the total, OIM, and DLM bias. Total bias is approximately 30 percent over this period. OIM bias is somewhat less than the total, and DLM bias is much greater. For both OIM and DLM, this indicates a systematic over-forecasting over the two years.

Figure 2.3. Bias—Total, Organization and Immediate Maintenance, and Depot-Level Maintenance



SOURCE: Produced using data from USAF, 2020–2021.

These results might seem somewhat surprising. As mentioned previously, DLM demands depend on PDMs and EOHs, both scheduled maintenance activities. However, forecast accuracy for those activities is much lower than for OIM activities.

DLM Forecast Error

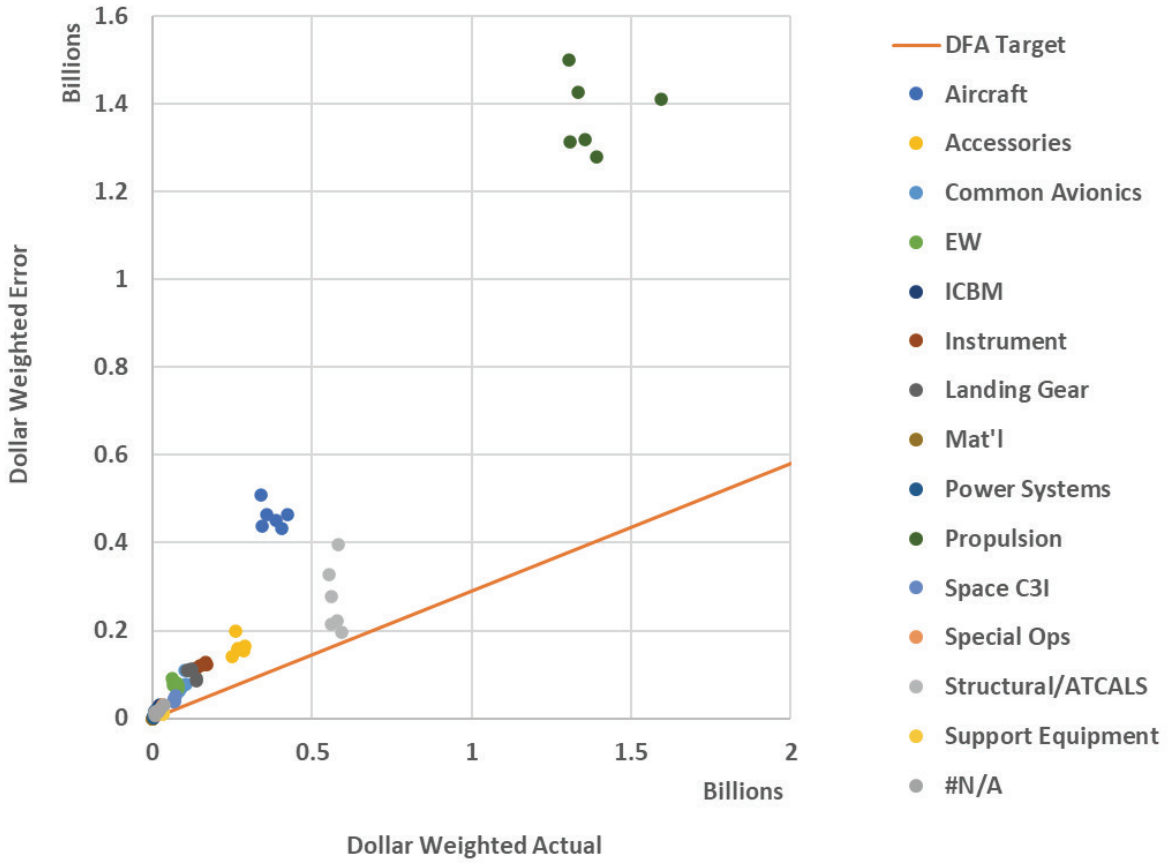
To better understand the drivers of this DLM forecast error, we examined the contribution of different part types to the overall error. The 448th SCMW comprises SCMSs that are each responsible for managing certain types of parts.[41] The forecast accuracy data provided by 448th SCMW included the SCMS that is responsible for each part. Thus, by showing error by squadron, we can see generally what type of part is driving total error. Figure 2.4 shows the total error versus total actual for each part type.[42] The error is shown on the *y*-axis and is the numerator of the DFA equation shown previously—the absolute value of the predicted demand minus the actual demand, weighted by dollar value. The *x*-axis is the denominator of the DFA equation—the total value of the actual demand. The orange line shows the 71 percent DFA (or 29 percent error) target. For example,

[41] See U.S. Air Force, 72nd Air Base Wing Public Affairs, *Tinker Air Force Base 80th Anniversary Units and Mission Structure*, 2022.

[42] Throughout this report, the term *actual* often stands alone. As related to forecasting the *expected* demand for spare parts required by USAF, the term *actual* refers to the number of parts that were, in fact, requested by USAF over a given period. For example, expected demand (forecasted) compared with actual demand (requested). The term *actual* is also used in comparing forecasted value of parts versus the value of parts requested in retrospect.

if the actual value of DLM repair parts in a given year is $1 billion, we would expect $290 million in error for a DFA of 71 percent (1–[$0.29B/$1B] = 71 percent).[43] This means that points above the DFA target line have more than 29 percent error, and the distance from the target line allows easy inspection of key drivers. This clearly shows that propulsion parts are driving DLM forecast error.

Figure 2.4. Depot-Level Maintenance Dollar Weighted Error Versus Actual by Part Type



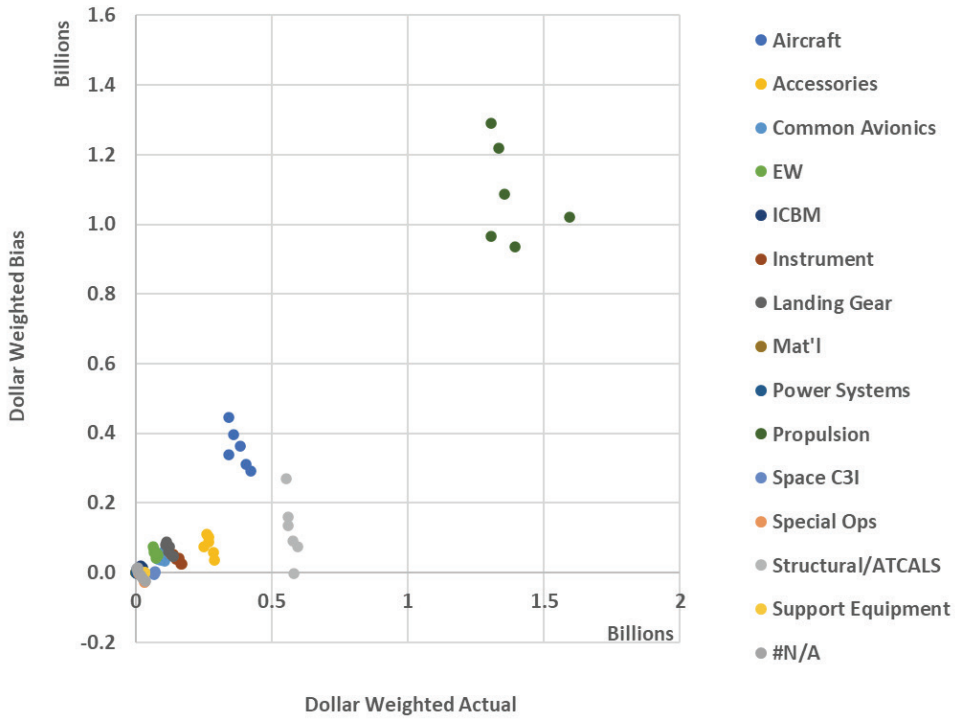
SOURCE: Produced using data from USAF, Forecast Analysis Comparison Tool, database, 2020–2021.
NOTE: EW = electronic warfare; ICBM = intercontinental ballistic missile; Mat'l = materiel; ATCALS = air traffic control and landing systems; #N/A = associated with an uncategorized item.

Figure 2.5 shows a similar chart but with bias shown on the y-axis. Values greater than zero indicate a positive bias resulting from over-forecasting. Clearly, the forecast error associated with propulsion parts is driven by systematic over-forecasting. Discussions with 448th SCMW personnel indicated that this systematic over-forecasting is a well-known phenomenon. EOHs usually have very

[43] The total value of the error and actuals mentioned here and throughout the remainder of the report is based on calculations using the latest acquisition cost, per DoD guidance. However, most of these parts are repaired, and repair costs can be much lower than acquisition cost. The dollar value of the error should not be interpreted as money that could have been spent elsewhere directly. It only provides a relative metric to assess forecasts.

specific work packages, with parts that can have long lead times; thus, the depot could be forced to order the entire work package and then install new parts only as required.

Figure 2.5. Depot-Level Maintenance Dollar Weighted Bias Versus Actual by Part Type



SOURCE: Produced using USAF, 2020–2021.

Organization and Intermediate Maintenance Forecast Error

As we turn to OIM demand, which is primarily driven by flight line failures and scheduled and unscheduled maintenance, we note that a common issue in demand forecasting for spare parts is intermittent demand.[44] DoD defines the following four levels of demand intermittency:[45]

- limited: demand in less than 10 percent of historical demand periods
- uneven: demand in 10–60 percent of historical demand periods

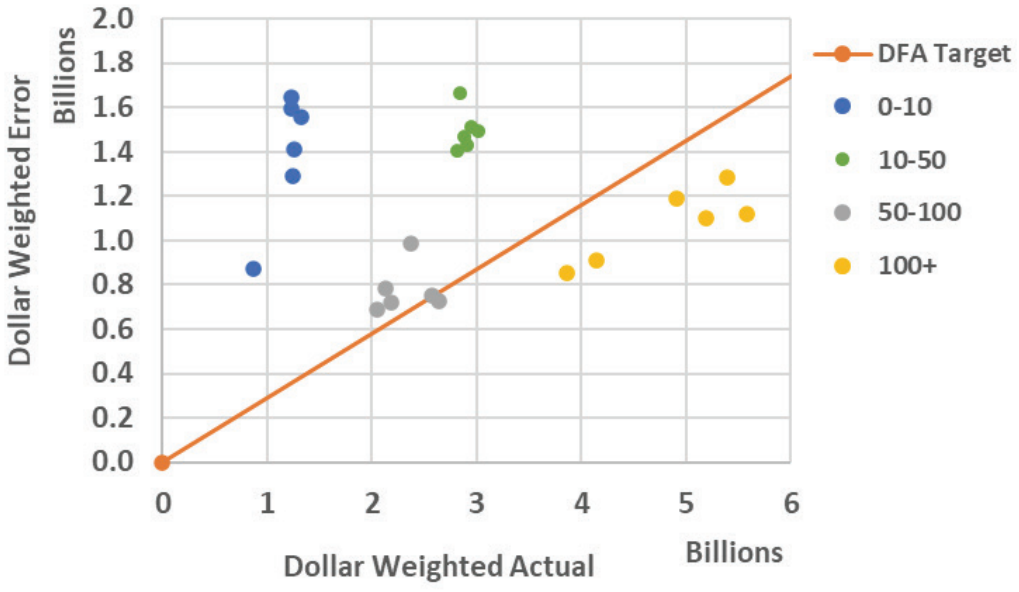
[44] Çerağ Pinçe, Laura Turrini, and Joern Meissner, "Intermittent Demand Forecasting for Spare Parts: A Critical Review," *Omega*, Vol. 105, No. 1, July 2021.

[45] Department of Defense Manual 4140.01 Volume 2, 2018.

- erratic: demand in 60–85 percent of historical demand periods
- continuous: demand in more than 85 percent of demand periods.

Forecasting demand becomes increasingly difficult as demands decrease to uneven or limited intermittency. There are limited ways to forecast uneven or limited intermittent demand levels, and we sought to (1) understand the magnitude of this effect and (2) investigate potential issues beyond demand intermittency. Unfortunately, the forecast accuracy data available to us included demands aggregated annually, so we could not specifically analyze intermittency. Instead, to provide a proxy for demand intermittency, we grouped parts according to the number of demands in a year to separate the effects of low-demand parts (ten or fewer demands in a year, a number that has been suggested in previous research)[46] and parts with larger annual demands. Figure 2.6 shows the dollar weighted error versus dollar weighted actual for OIM parts in a similar format as the previous figures for DLM parts but grouped by annual demand. The orange line again represents the DFA target of 71 percent (error of 29 percent). It shows that, for parts with an annual demand of 50 or more, forecasts are generally at or above 71 percent, while parts with demand less than 50 have error greater than 29 percent.

Figure 2.6. Organization and Immediate Maintenance Dollar Weighted Error Versus Actual by Demand Quantity



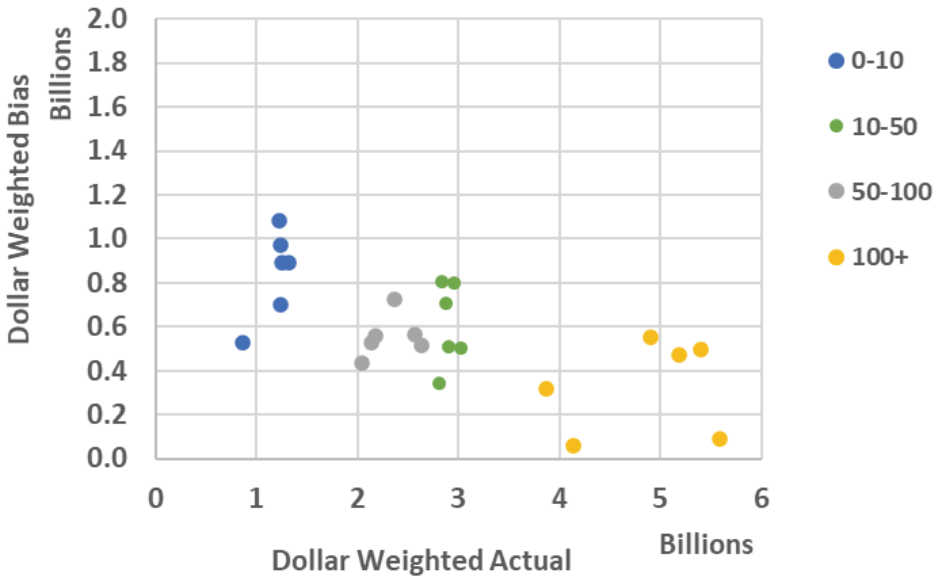
SOURCE: Produced using USAF, 2020–2021.

Figure 2.7 shows the bias associated with each demand quantity. All categories have a positive bias, indicating over-forecast, but the effect is larger for the low-demand items. These results were

[46] Mary E. Chenoweth, Jeremy Arkes, and Nancy Y. Moore, *Best Practices in Developing Proactive Supply Strategies for Air Force Low-Demand Service Parts*, RAND Corporation, MG-858-AF, 2010.

hardly surprising, and discussions with personnel from the 448th SCMW indicated clear awareness of these challenges. In the legacy system, inventory levels for these types of parts were set using a Logistics Management Institute–developed tool called Peak Policy and Next Gen (PNG).[47]

Figure 2.7. Organization and Immediate Maintenance Dollar Weighted Bias Versus Actual by Demand Quantity



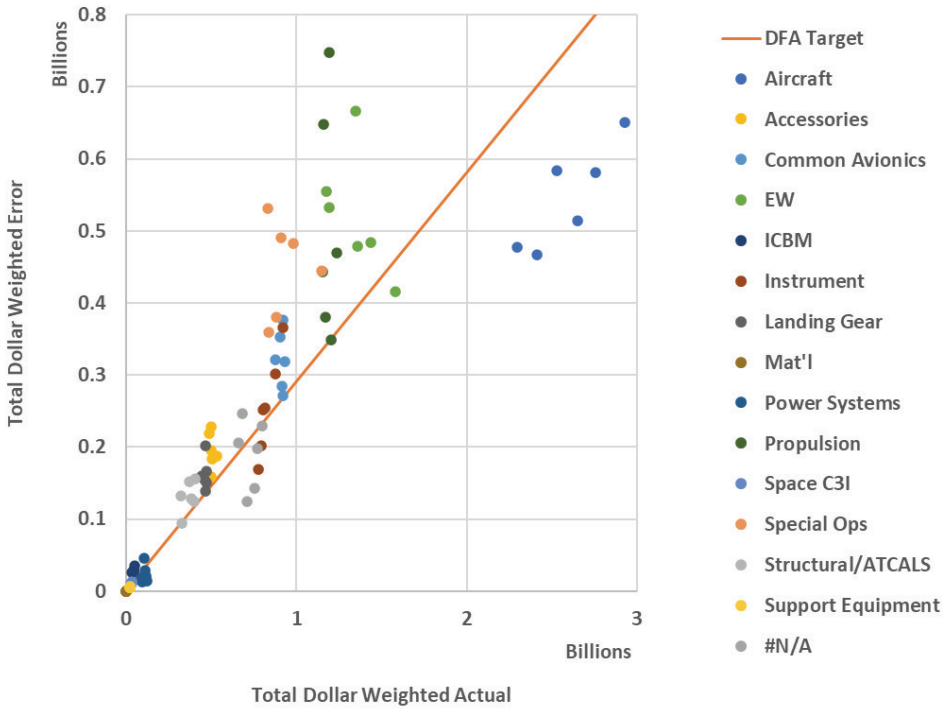
SOURCE: Produced using USAF, 2020–2021.

To further investigate the drivers of OIM forecast error beyond issues associated with low demand, we excluded all parts in the low-demand category (ten annual demands or fewer) and again plotted the total error versus actual (Figure 2.8). Forecast accuracy for most part types was near the DFA target with a few exceptions. Aircraft-related parts are forecasted with accuracy exceeding the target of 71 percent. However, a few part types are forecasted significantly worse in some years, including propulsion, EW, and special operations parts.

Discussions with personnel at the 448th SCMW indicated that many propulsion parts are *life limited*, meaning they are replaced prior to failure, and the Air Force Life Cycle Management Center (AFLCMC) produces forecasts for these parts. Over-forecasting often results from these replacements because they are planned but do not actually occur. For the special operations class of parts, 448th SCMW stated that a higher variance in mission profile (i.e., annual flying hours) results in more uncertainty in part failures and, thus, lower forecast accuracy. Although these explanations seem reasonable, we did not independently verify them.

[47] Tovey C. Bachman, Pamela J. Williams, Kristen M. Cheman, Jeffrey Curtis, and Robert Carroll, "PNG: Effective Inventory Control for Items with Highly Variable Demand," *Interfaces*, Vol. 46, No. 1, 2015.

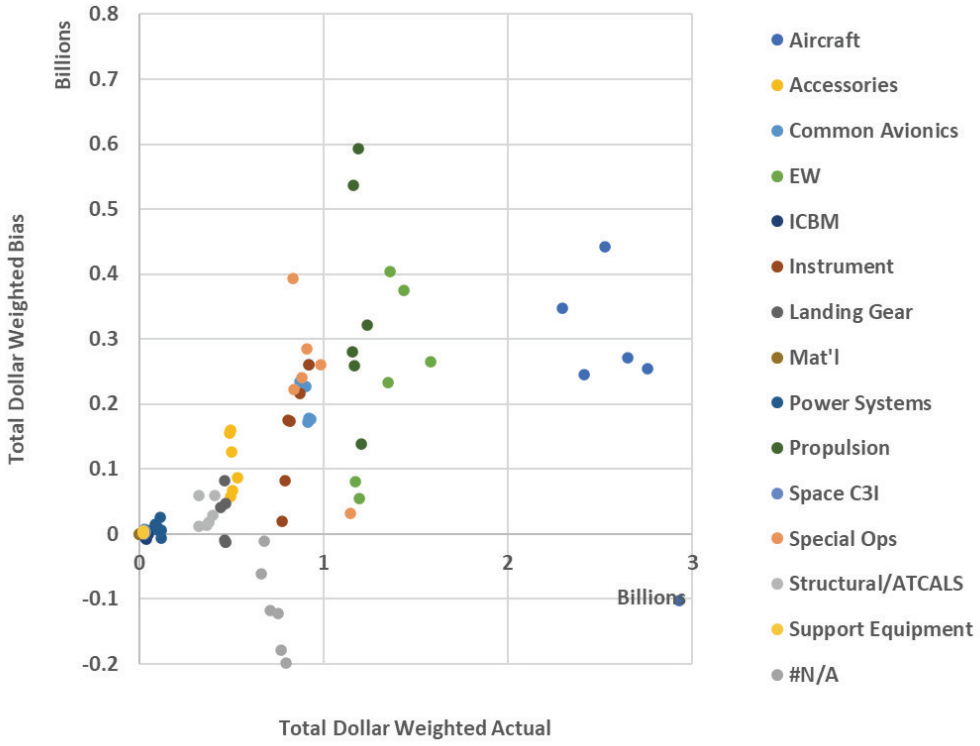
Figure 2.8. Organization and Immediate Maintenance Dollar Weighted Error Versus Actual by Part Type, Excluding Low-Demand Parts



SOURCE: Produced using USAF, 2020–2021.

Figure 2.9 shows the corresponding bias. Again, a positive bias is observed across most part types, indicating systematic over-forecasting.

Figure 2.9. Organization and Immediate Maintenance Dollar Weighted Bias Versus Actual by Part Type, Excluding Low-Demand Parts



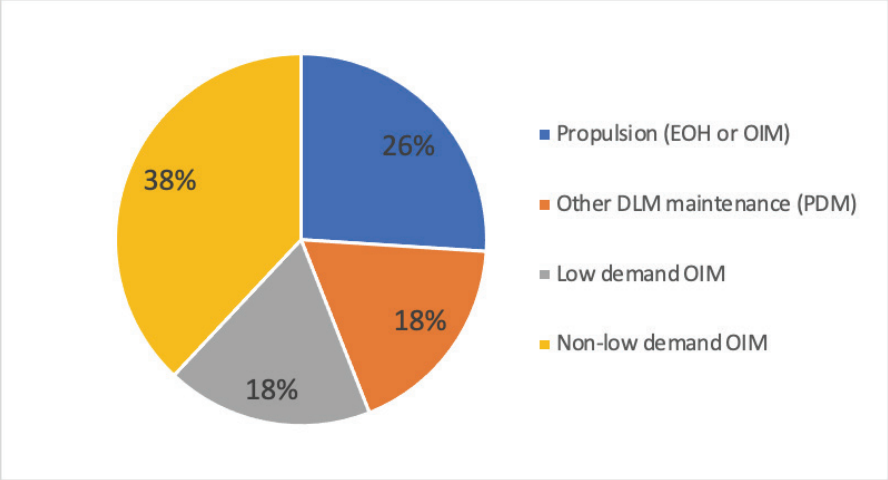
SOURCE: Produced using USAF, 2020–2021.

Summary for Depot-Level Maintenance and Organization and Immediate Maintenance

To summarize, our analysis highlighted the following sources of forecast error (see Figure 2.10):

- Error associated with propulsion parts (EOH or OIM) accounted for an average of 26 percent of forecast error across the periods analyzed.
- Error associated with other DLM maintenance (PDM) accounted for an average of 18 percent of forecast error.
- Error associated with low-demand OIM parts accounted for 18 percent of forecast error.
- Error associated with non-low-demand OIM parts accounted for 38 percent of forecast error.

Figure 2.10. Summary of Forecast Error Sources



SOURCE: Produced using USAF, 2020–2021.

Addressing each of these sources of error likely requires a different approach. USAF is in the midst of a major change to its supply chain planning system, which could address some of these issues. In the next section, we briefly describe the new system and qualitatively assess how it might address forecast error. In the final sections of this chapter, we review recent demand forecasting literature to identify potential improvements to demand forecasting and discuss the role of demand forecasting in supply chain management.

USAF Migration to Enterprise Supply Chain Analysis, Planning, and Execution

USAF has been using D200A as the primary planning component for supply chain management of spare parts for decades. Although the program has met the needs of USAF, the program's architecture of specialized systems and varied databases requires a high degree of manual manipulation, has become more difficult to maintain, and is difficult to link with enterprise-level processes. At the same time, commercial enterprise resource planning systems (ERPs) have become widely used throughout industry, by the other services, and by DLA.[48]

USAF had been considering moving toward an ERP since at least 2010, when it sponsored a demonstration of a commercial-off-the-shelf (COTS) tool for supply chain planning. In 2016, USAF awarded a contract for delivery of an integrated supply chain planning and scheduling service to ultimately replace D200A.[49] That effort became known as the Enterprise Supply Chain Analysis, Planning, and Execution (ESCAPE) program, which aims to improve USAF's supply chain planning capabilities across demand forecasting, supply planning, and inventory control.

[48] Accenture Federal Services, *Sustainment Lifecycle Phase Forecasting and the Impact on Business Outcomes*, July 2013, Not available to the general public.

[49] Sandy Windsor, "Escaping Today's Supply Chain Challenges," Air Force Sustainment Center, September 27, 2017.

The ESCAPE program office is implementing PTC's service parts management (SPM) software solution, identified by Accenture as a best-in-class solution for large, complex supply chains and in use today by such companies as Airbus, Boeing, Embraer, and Quantas.[50] ESCAPE consists of five mission area capabilities—demand planning, inventory planning, supply planning, exception management, and performance management—that together should achieve such benefits as improved DFA, reduced lead times, improved cost estimates, and reduced planning workload.[51] Of particular relevance is the way in which the demand planning capability could achieve improved demand forecasting. The demand planning capability in ESCAPE includes multiple demand forecasting techniques, where *demand* is defined as a request for a reparable or consumable part used during the repair and maintenance of USAF weapon systems, reparable end items, and equipment.[52]

ESCAPE will provide two types of forecasting techniques for OIM demand: statistical and causal. The specific statistical techniques available as part of the SPM product are proprietary but generally include most standard time series forecasting methods that forecast future demand based on historical data.[53] SPM includes a best-fit function that compares forecast methods and selects the best forecast for each part type. SPM also includes built-in functionality to calculate demand intermittence and variance and applies special forecast methods to this subset of parts.

Causal forecasting is used when demand is driven by use. In its initial configuration, causal forecasting will implement the legacy forecasting capability as described in the previous section, "Legacy Demand Forecasting," where causal types include flying hours, squadrons, equipment months, sorties, drone recoveries, and ammunition expenditures, and only one causal type is assigned. SPM provides the capability to assign multiple causal types but will be limited because of current USAF data limitations.

DLM demand is forecasted using the scheduled event maintenance (SEM) method. In this method, a bill of material and use rates of component parts are used to create the demand forecast. As in the legacy system, this method will be used for PDMs, EOHs, and NHA MISTR. Such event data as event (e.g., EOH), event product (e.g., F100 engine), event schedule (e.g., two overhauls per month), and event location (e.g., depot) will be loaded manually by users. SEM uses the line scrap rate, the replacement repair rate, and the line repair rate in conjunction with the event data to calculate the demand forecast.

USAF's implementation of SPM has been ongoing since 2016, and FY 2022 (the year this study was conducted) was the first year in which SPM was being used to generate demand forecasts. Discussions with personnel at 448th SCMW indicated that they were taking a measured approach in the rollout of ESCAPE, carefully managing change in the shift from D200A. In its first year of operation, ESCAPE was using the same input data as used previously by D200A but applying the best-fit statistical models for roughly two-thirds of the parts, which included the special forecast

[50] Blumberg Advisory Group, *Spare Parts Management Software State of the Art Benchmark Evaluation*, 2020. For PTC case studies, see Parametric Technology Corporation, "Case Studies," webpage, undated.

[51] 420th SCMS personnel, discussion with authors, January 3, 2021.

[52] Department of the Air Force Guidance Memorandum 2021-01 to Department of the Air Force Manual 23-122, *Materiel Management Procedures*, July 7, 2021.

[53] For a review of time series forecasting, see Jan G. de Gooijer and Rob J. Hyndman, "25 Years of Time Series Forecasting," *International Journal of Forecasting*, Vol. 22, No. 3, 2006.

models for intermittent demand mentioned previously. The remaining one-third continued to be forecast via the legacy straight-line causal forecast based on flying hours. The parts for which causal flying hour forecasts are used were determined by a manual assessment of the amount of change over time expected for certain parts, given program information.

For FY 2022, the 448th SCMW tracked and compared metrics for accuracy based on forecasts by both SPM and D200A. For FY 2022, the SPM forecast accuracy was higher than the D200A forecast accuracy for the same period (SPM 49 percent versus D200 42 percent). After FY 2022 the 448th SCMW no longer tracked D200A forecasts, so further comparisons are not possible. However, the SCMW anticipates additional forecast improvements as it continues to shift parts from the legacy forecasting techniques to the SPM statistical methods. It plans to officially evaluate additional improvements at the end of FY 2024.[54]

In the longer term, the 448th SCMW envisioned that the new demand planning process would more tightly integrate with the other supply chain planning functions and that this transformed process would place emphasis on combining the mathematical (data-driven) forecasting methods with improved collaborative processes to determine best possible demand forecasts.

Although there are some early signs of success, several challenges remain for the ESCAPE program. ESCAPE will still rely on the complex and disparate legacy systems and data sources, as it is not set to replace all of RMS. This has several implications. First, it means that although improved algorithms might provide modest improvements in forecast accuracy, there is limited ability to fully leverage the capabilities of ESCAPE by using data in new or better ways (e.g., by incorporating models that include multiple causal types) unless data are collected and integrated in different ways. Second, RMS is funded to do quarterly or monthly updates of key information sets, so forecast adjustments will remain periodic rather than be made in real time.

Insights from Academic Literature

Demand forecasting for spare parts is not a challenge unique to USAF. Companies across a variety of industries maintain inventories of spare parts to ensure equipment availability. Although specific goals might vary between military and commercial contexts, they share the primary challenge of balancing significant inventory purchase and holding cost with equipment downtime. Spare parts demand forecasting is a central component to this cost-availability trade-off and has therefore received significant academic and practical attention over the past 50 years that could help inform USAF going forward.[55] Several recent papers have provided comprehensive reviews of demand forecasting; we do not provide an additional comprehensive review here. Instead, we focus on specific findings highlighted in past research that might be of particular interest to USAF.[56]

---

[54] 420th SCMS personnel, email to authors, January 5, 2024.

[55] Pinçe, Turrini, and Meissner, 2021.

[56] For a comprehensive review of demand forecasting methods, see Pinçe, Turrini, and Meissner, 2021; and Boylan and Syntetos, 2009. For a review of the gaps between forecasting theory and practice, see Aris A. Syntetos, Mohamed Zied Babai, John Boylan, Stephan Kolassa, and Konstantinos Nikolopoulos, "Supply Chain Forecasting: Theory, Practice, Their Gap and the Future," *European Journal of Operational Research*, Vol. 252, No. 1, November 2015; and Andrea Bacchetti and Nicola

Think of Forecasting as a System, Not a Technique

Boylan and Syntetos (2009) introduced a framework for forecasting that is composed of a pre-processing phase, a processing phase, and a post-processing phase, which are all underpinned by a forecast support system.[57]

Pre-processing encompasses the classification of demand as fast or slow moving, intermittent or lumpy. Best practices for classification tend to segment demand by mean time between demands and the coefficient of variation of demand sizes.[58] Classification in this manner facilitates the selection of the forecasting method.

The *processing phase* is the application of a forecasting technique to generate the forecasted demand. The best forecast method varies based on demand pattern, and best practices typically suggest applying different methods specific to part type. Many forecasting techniques for spare part demands have been put forward and are typically categorized as either time series methods or causal (sometimes referred to as *installed base*) methods. A short review of various modeling techniques is provided in the next sections.

Finally, the *post-processing phase* includes adjustments to the statistical forecast made by the user. Such adjustments are common in practice, though implications for forecast accuracy have shown mixed results. Strategies for improvements in the post-processing phase include mechanisms to help users better understand the implications in adjustments to the statistical forecasts.

Traditional Models Continue to Provide Value

Time series models use historical data to generate demand forecasts, absent any contextual information. Classical methods (such as moving averages and exponential smoothing)[59] are the most widely used in practice, and there is empirical support for the continued use of these simple techniques for fast-moving demand.[60] For intermittent demand, several more-sophisticated models have been developed, including Croston's method and several subsequent methods that are variations of it.[61] These models forecast the interval between demand arrivals and the demand size separately via exponential smoothing, which tends to provide better forecasts for intermittent demand patterns. All these parametric time series methods require an assumption about the underlying demand distribution. Nonparametric time series models reconstruct the empirical distribution of the demand.

Saccani, "Spare Parts Classification and Demand Forecasting for Stock Control: Investigating the Gap Between Research and Practice," *Omega*, Vol. 40, No. 6, December 2012.

[57] John E. Boylan and Aris A. Syntetos, "Spare Parts Management: A Review of Forecasting Research and Extensions," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 21, No. 3, November 12, 2009.

[58] Aris A. Syntetos, John E. Boylan, and J. D. Croston, "On the Categorization of Demand Patterns," *Journal of the Operational Research Society*, Vol. 56, No. 5, August 25, 2004.

[59] A moving average essentially assigns equal weight to all past data points included in the forecast calculation, while exponential smoothing decreases the weight exponentially for points further back in time.

[60] Spyros Makridakis and Michèle Hibon, "The M3-Competition: Results, Conclusions and Implications," *International Journal of Forecasting*, Vol. 16, No. 4, 2000.

[61] J. D. Croston, "Forecasting and Stock Control for Intermittent Demands," *Operational Research Quarterly*, Vol. 23, No. 3, September 1972.

Traditional nonparametric models include bootstrapping methods that involve consecutive sampling from the data to construct an empirical distribution, though these methods have received considerable criticism and are not widely used in practice. Despite the advances in spare parts demand forecasting, there is no conclusive best forecasting method because different methods have shown superior results in different studies depending on the accuracy metrics used and characteristics of the demand.[62]

Emerging Models Are Showing Promise but Require Further Investigation

More-recent nonparametric methods include the use of ML techniques (neural networks [NNs] in particular) to learn demand patterns directly from the data through supervised learning.[63] These methods are particularly appealing (at least in theory) for predicting intermittent demand, given their nonlinear learning function. However, results to date have been somewhat mixed. In one case study on business aviation parts, researchers found that NNs showed improvements over traditional forecasting methods, provided that the model included sufficient demand features.[64] A different case, still related to commercial aviation, found that NNs outperformed some traditional forecasting methods but not others.[65]

There has also been renewed interest in causal models. These models attempt to account for the underlying demand-generating factors associated with the product's installed base, or the number of products still in use. The previous time series methods all rely on historical demands and, therefore, provide only reactive forecasts. Characterizing the factors that generate the demand allows proactive forecasts that anticipate future demand based on the expected changes to these factors. A literature review of causal forecasting for spare parts found the following three main sources of information that drive demand:

- the size and status of the installed base
- the maintenance policy
- the environmental factors that affect reliability.[66]

The review found that causal models in the literature are increasing in complexity over time through use of increased installed base information and that results show promising performance in terms of forecast accuracy, although evaluations and validations are not always conducted in standardized formats, making specific improvements difficult to assess.[67] Research being conducted for the

[62] Bacchetti and Saccani, 2012.

[63] M. Z. Babai, A. Tsadiras, and C. Papadopoulos, "On the Empirical Performance of Some New Neural Network Methods for Forecasting Intermittent Demand," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 31, No. 3, July 2020.

[64] K. Nemati Amirkolaii, A. Baboli, M. K. Shahzad, and R. Tonadre, "Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by Using Artificial Intelligence (AI)," *International Federation of Automatic Control-PapersOnLine*, Vol. 50, No. 1, July 2017.

[65] Babai, Tsadiras, and Papadopoulos, 2020.

[66] Sarah Van der Auweraer, Robert N. Boute, and Aris A. Syntetos, "Forecasting Spare Part Demand with Installed Base Information: A Review," *International Journal of Forecasting*, Vol. 35, No. 1, 2019.

[67] Van der Auweraer, Boute, and Syntetos, 2019.

Republic of Korea's military has suggested that incorporating additional feature sets related to the installed base (e.g., reliability and operating environment) can improve forecast accuracy.[68] A recent study by LMI found that data sparsity limits the potential of ML techniques to improve demand forecasting when using part demands alone, but ML models of maintenance events did have the potential to improve forecasts over traditional time series models.[69] In addition, recent RAND research suggests that NNs outperform traditional techniques for forecasting failures of certain parts.[70]

Data Integration and Cleaning Is a Challenge

As discussed, an area of recent active research is the expansion of causal models to predict spare parts demand driven by the increased availability, storage, and processing of data. Of particular interest is the link between part demand, use factors, and maintenance policies. As noted in one of the reviews, existing "systems typically spread these data over many disparate tables and/or separate systems. One apparently banal but nevertheless important challenge for software providers thus is to access and combine all relevant data."[71] A second paper, specific to the use of maintenance information in conjunction with part failure data to generate better forecasts, noted that:

> The main challenge to implement our method is the collection of accurate data. Our method requires keeping track of historical machine sales and discards to monitor the evolution of the installed base over time, as well as a history of past part failures, and information on (past and future) preventive maintenance interventions. Therefore, forecasting becomes the outcome of an inter-organizational process, where cooperation and information sharing (for example on historical failures, maintenance actions, and product sales) between different departments within the same company is needed.[72]

Demand Forecasting in the Broader Context of Supply Chain Planning

It is generally asserted that improvements to demand forecasting will reduce excess inventory or improve aircraft availability, but this assertion is difficult to demonstrate. The discussion on over- or under-forecasting the demand for spare parts is nuanced. Under-forecasting demand could result in

[68] Boram Choi and Jong Hwan Suh, "Forecasting Spare Parts Demand of Military Aircraft: Comparisons of Data Mining Techniques and Managerial Features from the Case of South Korea," *Sustainability*, Vol. 12, No. 15, July 2020; Hanjun Lee and Jaedong Kim, "A Predictive Model for Forecasting Spare Parts Demand in Military Logistics," *2018 Institute of Electrical and Electronic Engineers International Conference on Industrial Engineering and Engineering Management (IEEM)*, December 2018.

[69] Sergio Posadas, Carl M. Kruger, Catherine M. Beazley, Russell S. Salley, John A. Stephenson, Esther C. Thron, and Justin D. Ward, "Forecasting Parts Demand Using Service Data and Machine Learning," Logistics Management Institute, January 2020.

[70] Li Ang Zhang, Yusuf Ashpari, and Anthony Jacques, *Understanding the Limits of Artificial Intelligence for Warfighters: Volume 3, Predictive Maintenance*, RAND Corporation, RR-A1722-3, 2024.

[71] Syntetos et al., 2015, p. 16.

[72] Van der Auweraer, Boute, and Syntetos, 2019, p. 148.

parts not being available when needed to repair an aircraft, and thereby contribute to aircraft downtime. In that regard, over-forecasting demand is preferred, and generally is not problematic for low-cost parts. However, for expensive parts, over-forecasting demand results in increased "value of secondary item inventory," a supply chain metric where lower is better.[73] The relationship certainly makes logical sense, but demand forecasting is just one component of a complex planning process to ensure that the right part is at the right place at the right time.

To understand how demand forecasting fits into the broader context of supply chain planning, we conducted a thorough review of all metrics laid out in the *Supply Chain Metrics Guide*.[74] The metrics guide serves as a reference for the comprehensive set of DoD supply chain metrics that enables DoD to monitor supply chain performance. The guide includes diagnostic, functional, and outcome metrics across various attributes of the supply chain, including materiel readiness, responsiveness, reliability, cost, and planning. The guide provides a definition for each metric and describes the links between individual metrics.

One thing that the metrics guide does not include is an integrated view of all metrics that shows their interrelationships. Combining the defined logical linkages for each individual metric produces the integrated metrics diagram. An excerpt of the complete diagram, highlighted in Figure 2.11, shows that DFA is a logical driver of excess on-hand and aircraft downtime via wholesale supply availability, which affects not mission capable for supply (NMCS) backorders and, subsequently, NMCS. However, in each case it is one of six or more potential drivers. DFA is also a driver of excess inventory, which drives the value of secondary item inventory, which is another contributing factor to wholesale supply availability. The point of this exercise was to assess how improving DFA contributes to the overall outcome metrics associated with costs and materiel readiness. Although the logical connection can be made, the degree of impact cannot be assessed without more analysis, which was beyond the scope of this effort.

In its review of DFA, LMI was not able to correlate forecast error with either excess or shortfalls using historical data and concluded that excess and shortfalls are driven by a combination of factors.[75] A subsequent paper by Accenture provided simulated outcomes that did show that improved forecasting led to a small decrease in inventory and a more significant reduction in backorders.[76]

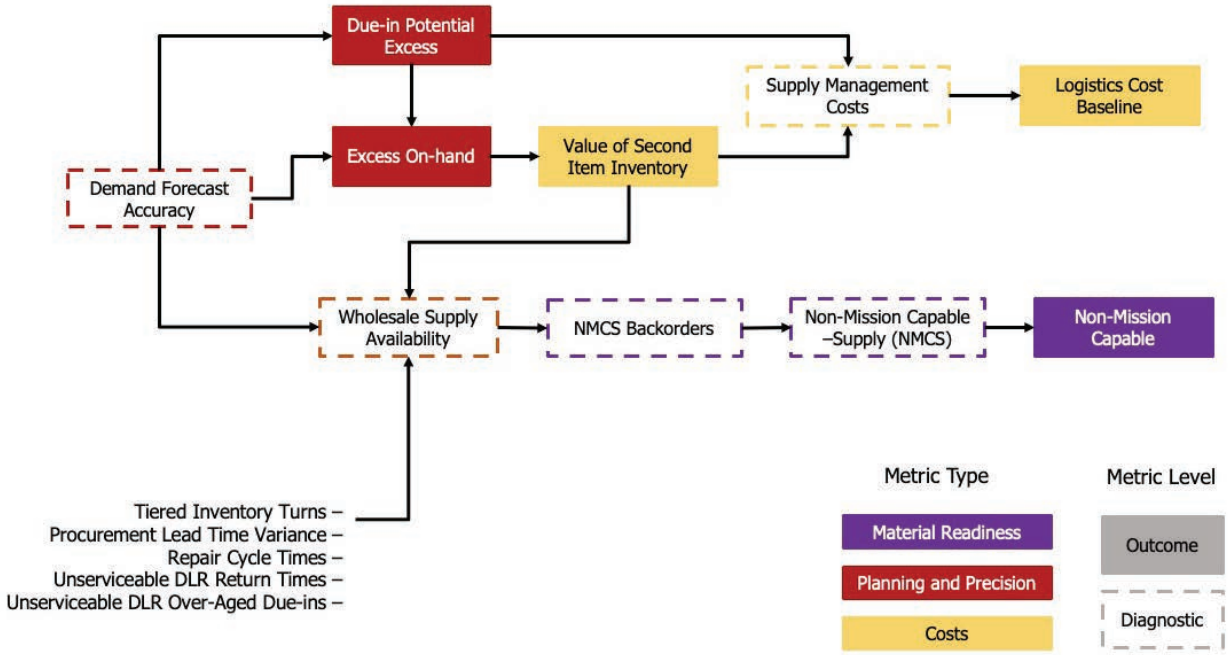
[73] DoD, 2021, p. 52.

[74] DoD, 2021.

[75] Walter D. Atchley, Dorothy M. Clark, Salvatore J. Culosi, Lori Dunch, Robert C. Kline, Thomas E. Lang, Randy L. Moller, Matthew R. Peterson, and Michael R. Pouy, *Lifecycle Forecasting Improvement: Causative Research and Item Introduction Phase,* Logistics Management Institute, Report DL920T1, November 2010.

[76] Accenture Federal Services, 2013.

Figure 2.11. Integrated Representation of Demand Forecast Accuracy–Related Supply Chain Metrics



SOURCE: Produced using information from DoD, 2021.
NOTE: DLR = depot-level reparable.

Observations and Findings

In summary, our analysis led to the following observations and findings:

- Demand forecasting for spare parts is a topic that has received significant attention both inside and outside USAF for decades. There are a variety of methods to forecast spare parts demand, though there is not a one-size-fits-all approach that has been shown to be the best in all cases. The best approach tends to depend on such things as characteristics of the demand and availability and quality of the data.
- Because there is not a single solution to demand forecasting, we examined the sources of demand forecast error based on recent USAF forecast accuracy data that might highlight potential solutions. This analysis revealed sources of error; personnel within the 448th SCMW who study the problem were mostly aware of these errors, but they might not be commonly understood across the broader logistics, engineering, and force protection (A4) community. Most notably:
  - Demand for parts associated with programmed actions (e.g., EOHs, PDM) is a large source of forecast error.
  - Parts with low demand contribute disproportionately to forecast error.

- USAF is in the midst of a major change to the way it forecasts spare parts demand as it shifts from D200A to ESCAPE. ESCAPE will use a best-in-class software product that could address several sources of the forecast errors identified and is already showing promising results. In particular, the best-fit time series methods could improve forecast accuracy for high-demand parts, and the special methods for intermittent or highly variable demand could improve forecast accuracy for low-demand parts, as defined in this report.
- There is ongoing research both within DoD and across industry on the use of more sophisticated techniques, such as AI/ML, for demand forecasting, though additional research is needed prior to broad implementation for USAF. These techniques could be the most beneficial to improve causal models where current time series forecasting techniques do not perform well. These methods will likely require better integration of data, including part demand, use factors, and maintenance policies.
- Demand forecasting is just one part of supply chain planning, and it is unclear the degree to which errors in demand forecasting (versus performance of other elements of supply chain management) are resulting in increased aircraft downtime.

Chapter 3

Bots for the USAF Supply Chain

USAF service members, civilians, and contractors are often burdened with repetitive tasks that occupy time better spent elsewhere. To alleviate time spent on these tasks, free people to focus on analytical work, and increase efficiency and consistency of repetitive processes, USAF has endeavored to apply business process automation. Process automation presents a significant opportunity for USAF to expedite many of its high-volume, manual processes, including those required for supply chain management. To this end, the A4 community has already started working on *bots*, a type of software that can automate manual processes, scrape data, and even use AI/ML to aid decisionmaking. However, process automation has risks; adopting any new technology requires consideration of the potential benefits and drawbacks. Thus, HAF/A4 asked us to investigate how it should move forward with applying bots to the management of its internal supply chains. In this chapter, we seek to provide a deeper understanding of what bot-enabled increased automation would entail for USAF's supply chains, including the types of bots that add the most value, methods for measuring return on investment (ROI), the key parameters that must be defined before development, and the risks that automation presents.

Methodology

Our analysis of bots for USAF supply chains used a variety of qualitative methods. We used a significant literature review to define and characterize bots. We held numerous discussions with USAF stakeholders and automation and supply chain subject-matter experts (SMEs) inside and outside RAND to delineate the bot progress made so far by USAF and to identify ideas for further bot investments by the A4 community. Bots for USAF supply chains go beyond automation of routine tasks; they can be cross-functional and even analytical.

Bots are a growing technology resource that can be applied to diverse uses. However, we also found that bots are not consistently defined in technology literature, and so confusion can arise as to exactly what bots are and how they can be applied. This combination can make it challenging for decisionmakers to figure out how to approach bot development or which tasks to prioritize. Therefore, we begin by introducing a comprehensive bot taxonomy developed by Lebeuf et al. (2019) to classify bot features (a more detailed description of the taxonomy appears in Appendix B).[77] These features are pivotal to understanding the benefits that a bot could provide and the security risks it could introduce. This trade-off between risk and reward is certainly not unique to bots, but as USAF

[77] Carlene Lebeuf, Alexey Zagalsky, Matthieu Foucault, and Margaret-Anne Storey, "Defining and Classifying Software Bots: A Faceted Taxonomy," *2019 Institute of Electrical and Electronic Engineers/Association for Computing Machinery 1st International Workshop on Bots in Software Engineering (BotSE)*, 2019.

increases its use of bots, it is essential that the service examine both. Thus, we describe the status of bot development and use in USAF and in commercial supply chains and discuss potential risks involved. Finally, we describe a potential bot concept that we believe will be valuable to merging maintenance and failure data for a variety of purposes, including demand forecasting.

Defining and Describing Bots

An initial challenge for bot development is the lack of consensus on what constitutes a bot. This can confuse or limit policymakers' ability to understand bot technologies and how to approach their development. In the case of robotic process automation as a service (RPAaaS), it can also lead to challenges in comparing bot service providers. In this section, we present a definition of *bot* and a review of a comprehensive bot taxonomy to characterize bots (with further details in Appendix B). We believe these should be useful to USAF as it goes forward with bot development and deployment.

From its inception, the term *bot* has been defined many ways. Simple definitions include "a software version of a mechanical robot"[78] and "a computer program that performs automatic repetitive tasks."[79] However, as bots have become more sophisticated and diverse, more conceptual characterizations have been proposed, such as "the bridge between data and action."[80]

One reason that the term *bot* has been difficult to define precisely is that numerous other software terms define similar technologies. For instance, a *script* is a small piece of software that performs a task but does not perform any significant computing on its own. Scripts are typically written in the eponymously termed *scripting language* (which is different from *programming languages* in that it is typically interpreted, not compiled). *Computer programs* are larger pieces of software, written in the aforementioned programming languages, that perform significant computing. A bot lies somewhere between these definitions, depending on the level of sophistication and the software used to create the bot.

In this report, we leverage the broad definition proposed by Lebeuf et al. because it robustly applies to all types of bots. It is "the interface that *provides the services to the user* . . . a bot is everything required to present the service to the user . . . and provides additional value on top of its services."[81]

Examples of services a bot can provide include automation, lowering the barrier of access to data, pulling from multiple data sources, and presenting visualizations. Software services are either external or internal to the bot, where internal services are local or offline and external services are online. Some bots can provide internal and external services. A bot can have a combination of internal and externally accessed services. For example, in the context of USAF, a bot might interact with data on such an enterprise system as the Logistics, Installations, and Mission Support—Enterprise View (LIMS-EV) (external) and generate a presentation of those data with a local PowerPoint document

[78] Andrew Leonard, *Bots: The Origin of New Species*, Penguin Books Limited, 1998.

[79] "Bot," Merriam-Webster, webpage, undated.

[80] Suman Deb Roy, "What Bots May Come: An In Depth Discussion of a Learning Architecture for the Next Paradigm," *Chatbots Magazine*, March 20, 2016.

[81] Lebeuf et al., 2019.

(internal). Users interact with software bots in multiple ways, including the command line, graphical user interfaces, touch interfaces, and spoken and written language.

With this working definition of a *bot*, we describe a structured way for USAF to approach bots in terms of function and level of sophistication. Then, we present a taxonomy that captures the primary attributes of a bot's environment, execution, and interactions with others. This taxonomy will allow USAF to clearly define a bot application prior to its development. Clear, *a priori* characterization of a bot holds a variety of advantages, including identification of potential security vulnerabilities, planning for bot development cycles, and development of a bot as user-friendly and useful as possible from the outset.

Characterizing Bots

A bot can be classified along two primary dimensions: its function and its level of sophistication. In this section, we present these dimensions as a starting point for USAF to characterize its current and future bot development.

As stated above, bots have been applied to automate processes by collecting, cleaning, analyzing, transmitting, and presenting information. It is important for USAF bot requesters and developers to consider the function they wish the bot to perform. Simply wanting a process to be automated is not sufficient. For instance, *crawler bots* might be developed to continuously scour data for misalignment between data systems, or they might continuously examine procurement data for orders that deviate from the norm as orders are introduced to the system. A *scraper bot* could be developed to pull cost and availability data from the websites of suppliers, such as DLA, and consolidate them into a cost database.

Moreover, *transactional bots* could be developed to place orders or regularly pull and consolidate data from systems that govern maintenance, supply, and operations. Then, a *productivity bot* could schedule exercises, and a *documentation bot* could create briefings to inform leadership of inventory levels, operational readiness, and upcoming exercises. Crawlers, scrapers, transactional bots, productivity bots, and documentation bots all exist to automate, but they function in different and important ways.

Table 3.1 provides a more comprehensive list of bot types and their corresponding functions, including such popular types as crawlers that capture data from websites, transaction bots that work on a user's behalf, and documentation bots that create reports and briefings. These different types can be instantiated using robotic process automation (RPA) or intelligent automation (IA).

Table 3.1. Bot Functions

Type	Function
Crawlers	Run continuously in the background to fetch and store data from websites, application programming interfaces, etc.
Scrapers	Used to download data from the web, sometimes to republish the data elsewhere
Information bots	Bring information to users

Type	Function
Transactional bots	Work on the user's behalf, interacting with external systems to automatically execute transactions when a condition is met
Productivity bots	Increase productivity by automating simple tasks
Collaboration bots	Enable users to communicate and collaborate
Hacker bots	Distribute malware and exploit vulnerabilities in security
Testing bots	Detect bugs and errors in code
Code bots	Facilitate more-efficient coding
Documentation bots	Produce documentation using input data
Translation bots	Translate data from one language to another

SOURCES: Features information from Lebeuf et al., 2019; Radware, "Types of Bots: An In-Depth Guide by Radware," undated.

As the applications of bots have expanded over time, the universe of bots has become too large for a function-based classification system to be fully sufficient by itself. Thus, it is additionally helpful to characterize by level of sophistication. In this report, we focused on two levels of sophistication: RPA and IA. The primary distinction between these two is that an RPA bot is explicitly defined by programming logic, while an IA bot uses ML algorithms. Note that both RPA and IA can be highly sophisticated pieces of code, but the former relies on the programmer to flesh out the entire capability, while bots constructed with IA are trained on huge datasets and are less prescriptive (and thus also are more likely to be black boxes). Although these categories might not be comprehensive of all bot types, they represent the two primary classes of bots described in technology literature and produced by commercial enterprise.

Although ML and AI techniques are being used in USAF, we are not aware of IA bot applications using these techniques that are currently in use. Challenges to IA bot development in the very near term include airmen's limited technological expertise and unsuitable data for IA applications. These challenges are discussed in more detail later in this chapter. Although these challenges may be overcome in the future, we focused the bulk of our attention on RPA bots and the current capabilities of the A4 community.

Robotic Process Automation

RPA mimics the behavior of a human interacting with software applications.[82] These types of bots are valuable for automating processes that

- are manual and repetitive (e.g., copying data from one data source into another)
- are rule-based (usually using "if-then-else" logic, where the programmer has no uncertainty of underlying algorithms)

[82] "What Is Robotic Process Automation?" webpage, Association for Intelligent Information Management, undated.

- involve readable, electronic, standardized inputs (e.g., printed text; if one wanted a computer to "read" handwritten text, that would require ML)
- involve high-volume processing (e.g., repetitive calculations on large datasets)
- have a low exception rate (e.g., a formula applies to most observations in a dataset).[83]

RPA bots are useful for speeding up tedious processes in environments where system change is infeasible or in progress. In USAF, the types of processes that are the best candidates for RPA bots include creating regular briefings, consolidating data from disparate systems, and inputting information from forms. These are processes that do not involve uncertainty, opinions, or judgment calls, but rather are tedious, routine tasks that put a heavy burden on humans. RPA also helps avoid errors that humans are prone to make—for example, when inputting data from one system into another—because RPA bots never deviate from their encoded routines, do not misread information as human eyes can, and do not make keystroke errors. However, bots can still commit errors stemming from bugs in programming or by amplifying data that were initially entered incorrectly. Because the bot will never deviate from its code, this can sometimes cause much greater damage than a human user would have caused by making a mistake only once. A human might also catch underlying data errors if a data point looks wrong.

RPA has proven useful in multiple areas of business, including supply chain management, human resources, and accounting. Table 3.2 lists some common business areas and processes in commercial industry in which RPA has been valuable. USAF performs all of these processes daily. For supply chain management, RPA has improved inventory management, freight management, demand planning, and other processes.[84]

Table 3.2. Common Processes Automated Using Robotic Process Automation

Business Area	Processes
Supply chain management	• inventory management • demand and supply planning • work order management • invoice and contract management • processing returns • freight management
Human resources	• data entry • payroll • time management • benefits • recruitment • compliance and reporting

[83] Bill Nystrom, *Democratizing Automation for Every Airman*, PowerPoint presentation, UiPath, undated.

[84] Nystrom, undated.

Business Area	Processes
Finance and accounting	• vendor management • procure-to-pay • collections • sales order management
IT services	• server and application monitoring • routine maintenance • batch processing • email processing and distribution • password reset and unlocking • backup and restoration

SOURCE: Features information from Nystrom, undated.

Planning to build a bot requires several criteria to be considered (Table 3.3).[85] To ascertain whether RPA is applicable, bot developers should consider process predictability and whether manual work is involved. Also, if systems are upgraded frequently, it is important to consider how a bot might break or become obsolete with regular system changes. To assess whether RPA will add value, developers should consider transaction volume and the number of systems that a user must access to complete the process. If a process is repeated often and involves navigating multiple systems, RPA might add significant efficiency.

Table 3.3. Criteria for Determining Utility of Robotic Process Automation for Bot Functions

Criterion	Description
Manual work involved	RPA is best applied to expediting manual processes that take up too much time. If a process involves repetitious, tedious manual work, RPA might be a great solution to free up workers' time.
Number of systems accessed	Bots can access and go between systems much faster than a human can. If the process involves accessing multiple systems, automation might be valuable. At the same time, increasing the number of systems involved also increases the complexity of the automation, as each system could require distinct credentials and connection protocols.
Transaction volume	Referring to how often a process is executed, transaction volume is an important criterion to consider. Processes that are not executed often might not be worth the investment.
Error or rework prevalence	RPA adds consistency to processes and prevents some human errors from occurring. If the process involves such error-prone operations as manually entering data, performing manual information queries, or making complex calculations using data from multiple systems, automation could eliminate several errors.
Process predictability	RPA works only for processes that are rule based and without uncertainty, where identical inputs will yield identical outputs.

[85] HAF/A4 Logistics Automation, "Technology Modernization Fund: Full Project Proposal," PowerPoint presentation, October 2021.

Criterion	Description
System upgrade frequency	System upgrades can affect how a bot functions, potentially causing malfunctions. If systems are continuously changing because of upgrades, the lifetime of a bot might not be as long as desired.

SOURCE: Features information from HAF/A4 Logistics Automation, 2021.

Before embarking on an automation effort, it is also important to fully document the workflow for the bot and identify sources of potential complexity. This will inform how difficult programming the bot will be and the financial and time investment that might be required. Table 3.4 includes important sources of bot complexity. Especially relevant for USAF are those indicators related to leveraging multiple applications and navigating disparate data systems using varying types of connectivity across multiple levels of security. Importantly, RPA bots should be viewed as a temporary means of connecting systems, whereas more-fundamental system integration is preferred where possible. However, integrating systems can be hard for many reasons, including incompatible software (e.g., Mac, Windows, or Linux), varying degrees of security (e.g., public information, controlled unclassified information, or classified information), and credentialed access (e.g., common access card [CAC], certain work users). For instance, integrating supply and maintenance systems might be more challenging if USAF supply personnel should access only supply systems and USAF maintenance personnel should access only maintenance systems.

Table 3.4. Indicators of Bot Complexity

Complexity Indicator	Description
Security	If a bot is designed to access secure systems, the bot must be given credentials or at least prompt the user. This adds complexity to the design and coding of the bot, especially if it goes between multiple systems that each carry different security criteria and privilege requirements. For example, each system may require a CAC or personal identity verification and distinct username and password. Alternatively, systems could be designed for different job types (e.g., supply versus maintenance), and a bot would need clearance from all parties to function.
Type of connectivity	A bot could leverage different types of connectivity, such as local networks, servers on a cloud, and internet websites. Each of these could require its own connection protocol, adding complexity to programming.
Number of applications	The more applications that a bot uses, the more complicated its workflow will become. Some applications require special user credentials, which the bot would need to access.
Number of keystroke steps	Each keystroke needs to be automated. Thus, the more keystrokes exist within a process, the more programming is required. However, one of the primary purposes of RPA is to reduce user keystrokes. So, although keystrokes complicate programming, a reduction in keystrokes could save user time.
Data structure	Retrieving and manipulating data from complex data structures or multiple distinct data structures could require complex programming.

Complexity Indicator	Description
Systems integration	Bots could act as a temporary means of bridging data systems. However, where possible, integrating systems rather than using a set of bots could provide the greatest efficiency returns.

SOURCE: Features information from HAF/A4 Logistics Automation, 2021.

When deciding whether automating a process will yield significant ROI, the value of an RPA automation can be measured by comparing the cost of implementing and maintaining an automation with the cost of continuing to perform the task manually. We augment a framework proposed by Gružauskas and Ragavan and propose an ROI that captures efficiency gained through automation, $Efficiency = EP - NP$, where EP is the existing process expense and NP is the new process expense.[86] The existing process expense is

$$EP = Total\ transaction\ volume \times Average\ processing\ time \times Average\ wage.$$

It is important that all parameters use the same unit of time (minutes, hours, etc.). The new process expense is given by

$$NP = Cost\ of\ bot\ development + Implementation + Maintenance.$$

To make EP and NP comparable, the length of time used for the *total transaction volume* must be long enough (at least one year). If $E > 0$, the automation will add efficiency (e.g., saving money or resulting in gained productivity).

Although this equation captures costs, it is also important to consider other criteria that might be more difficult to monetize. For example, it is difficult to monetize the value of refocused time, called *gained productivity*, that permits airmen to accomplish tasks that require deeper thought or judgment while an RPA bot handles manual tasks.

One of the stated goals for bot development in USAF is to enable more time to be spent on tasks that require more thought and judgment. Other metrics focus on the total cost savings from increased performance, how often the bot is used, or the value of error reduction.[87] Other helpful criteria to consider include the following:[88]

- **Process timing**: Automation can be leveraged to optimize a process. For example, an automation can be running in the background or during hours when humans are not working.
- **Service availability**: The percentage of time the service should have been available can be compared with the percentage of time it actually was available.

[86] Valentas Gružauskas and Diwakaran Ragavan, "Robotic Process Automation for Document Processing: A Case Study of a Logistics Service Provider," *Journal of Management*, Vol. 36, No. 2, December 2020.

[87] Bart Teodorczuk, "How to Measure RPA Success? A Guide to Robotic Process Automation Metrics," *Flobotics* blog, December 23, 2021.

[88] "Measuring RPA ROI—How to Do It Right?" *Digital Workforce* blog, July 1, 2020.

- **Personnel satisfaction**: Automations might improve productivity, which could improve morale and reduce staff turnover.

Once a bot has been implemented, it is important to measure the value being added. Fortunately, there are several indicators to inform how well an automation is performing,[89] eight of which are illustrated in Table 3.5. These indicators are meant to capture the value of how often the RPA is being used and the costs of downtime during maintenance. Leadership should develop documentation to regularly track these indicators after a bot is deployed.

Table 3.5. Automation Performance Metrics

Metric of Automation Performance	Description
Velocity	The average time it takes to perform an automated process; helps calculate time and cost savings
Utilization	How often the automated process is performed
Accuracy	Frequency with which the automated process is executed with errors
Break-fix cycles	Frequency with which an automated process malfunctions and requires maintenance
Break-fix person hours	The number of hours spent fixing a bot
Break root causes	The reasons that bots are breaking in the first place
Average automation uptime	How often a bot is available for tasking
Business value lost in downtime	Quantification of performance lost because of break-fix cycles by subtracting the value of downtime from the annual expected business value

SOURCE: Features information from "Measuring RPA ROI—How to Do It Right?" July 1, 2020.

Intelligent Automation

IA is more sophisticated than RPA, adding advanced technologies, such as AI, ML, and natural language processing (NLP).[90] IA bots are capable of tasks that RPA bots are not, including the following:

- analyzing patterns to inform strategic decisions

[89] Bruna Sofia Simoes, "Infographic: 10 Metrics You Should Be Tracking to Drive RPA Success," Blueprint, webpage, January 22, 2021.

[90] "What Is Intelligent Automation?" International Business Machines Cloud Education, webpage, undated; Pascal Bornet, "A Framework for Explaining the Power of Intelligent Automation," *Wevolver*, February 15, 2022.

- applying judgments to address uncertainty
- processing such nonstandardized inputs as varying formats and handwritten documents
- using NLP to understand logic within written language.

IA can improve accuracy in repetitive processes that require decisionmaking and reduce business costs by freeing up more personnel time for executing processes beyond repetitive tasks.[91] International Business Machines (IBM) promotes a "Supply Chain Intelligence Suite" to help businesses add resilience and agility by combining disparate systems, arguing that this will "lead to faster problem resolution and more efficient supply chain operations."[92] Current research on IA applied to supply chains is mainly focused on such themes as supplier risk management, decision models, and network design.[93] For example, if a supplier becomes unavailable because of disruptions, an IA bot might analyze historical data pertaining to costs, quality, and timeliness to choose the next supplier(s) faster than a human could.

ML has been implemented to enhance demand forecasts by applying data analysis to handle uncertainty; to the best of our knowledge, these applications have not used bots.[94] For this type of application, data must be trusted to reflect reality. Otherwise, the bot could predict a surge in demand when none comes to fruition, leaving the user with excess inventory. Alternatively, the bot will not predict an actual surge, parts will not be ordered, and weapon systems will be non–mission capable. It is important to remember that when it comes to handling uncertainty, there are no guarantees that the bot will make the correct decision. Rather, developers must quantify a level of confidence in results.

IA's initial use might be in helping USAF analysts process unstructured information that comes in unstandardized formats or includes handwriting. This incremental step above RPA could give USAF a chance to see what AI/ML can accomplish in a constrained environment. Such IA bots process large numbers of documents that come in varied forms, applying ML to identify the important information to extract and analyze.

There are challenges to implementing IA. First, IA generally requires sophisticated knowledge to implement, including advanced programming skills and a deep understanding of ML and data science. Bugs introduced during the programming and training of a model can lead to errors that are difficult to detect. In industry, AI bots have been applied to identify the best candidates from thousands of resumes; however, unfortunately, poor training data made these bots biased against certain types of people.[95] *Forbes* reports that a poorly trained bot "may promote biased hiring because of its reliance on unconsciously prejudiced selection patterns like language and demography."[96]

---

[91] "What Is Intelligent Automation?" undated.

[92] "What Is Intelligent Automation?" undated.

[93] Farheen Naz, Anil Kumar, Abhijit Majumdar, and Rohit Agrawal, "Is Artificial Intelligence an Enabler of Supply Chain Resiliency Post COVID-19? An Exploratory State-of-the-Art Review for Future Research," *Operations Management Research*, Vol. 15, Nos. 1–2, September 2021.

[94] See, for example, Babai, Tsadiras, and Papadopoulos, 2020; and Amirkolaii et al., 2017.

[95] Jeffrey Dastin, "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women," Reuters, October 10, 2018.

[96] Nish Parikh, "Understanding Bias in AI-Enabled Hiring," *Forbes*, October 14, 2021.

Another challenge is one associated with a common use of IA: conversational bots or chatbots that assist customers with IT support. These bots apply AI to understand the description of a problem from a customer and make judgments regarding the appropriate solution to the problem. If a bug exists in the programming or data that an IA bot is trained on, a conversational bot could give the wrong answer to a question. This would cause significant problems if the user were a mechanic asking for the correct part to use or a pilot asking for destination coordinates. Furthermore, a conversational bot might give the correct answer to a different question from the one it was asked. For example, if a user asked what type of part goes on an F-16, a bug-prone conversational bot could give the correct answer for an F-15 and not realize the error.

It is also important to be cautious about the decisions that are left to a bot. IA should be used to inform strategic decisions but should not be trusted to make these decisions by itself. After all, training an AI bot on bad data will teach it to make bad decisions.[97] Because of the dangers of putting confidence in results that stem from inaccurate data, one application of an automation, be it RPA or IA, is to look for inconsistencies in data. This type of application also adds cybersecurity, as malicious attacks often seek to alter data.

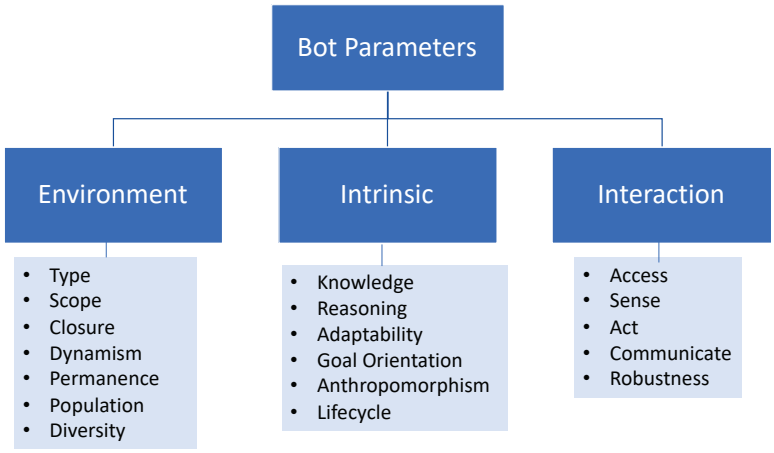
A Comprehensive Taxonomy to Characterize Bots

Before USAF can strategize its bot investments and begin bot development, it is imperative to understand and document the primary characteristics of a bot. Although most developers have a process for describing a bot prior to its development, a full taxonomy can raise questions for the bot's use that might not have been previously considered. Without a full taxonomy, HAF/A4 might have an incomplete picture of the impact a bot will have on systems and other users, which could cause conflicts with other systems or could introduce additional risks.[98] We propose using the comprehensive taxonomy published by Lebeuf et al. (2019), which characterizes a bot according to three primary dimensions: *environment*, *intrinsic* (how the bot functions), and *interaction*. Within each of these dimensions, numerous attributes describe the key information we need to know about a bot, illustrated in Figure 3.1. The full taxonomy is presented in detail in Appendix B.

[97] Tammy McClausland, "The Bad Data Problem," *Research-Technology Management*, Vol. 64, No. 1, 2021; RAND SMEs, guided discussions with the authors, August 2022; Stephen C. Slota, Kenneth R. Fleischmann, Sherri Greenberg, Nitin Verma, Brenna Cummings, Lan Li, and Chris Shenefiel, "Good Systems, Bad Data? Interpretations of AI Hype and Failures," *Proceedings of the Association for Information Science and Technology*, Vol. 58, No. 1, 2020.

[98] For more information on risks associated with bot development, see Chapter 4 of this report.

Figure 3.1. Primary Characteristics in a Comprehensive Bot Taxonomy



SOURCE: Adapted from Lebeuf et al., 2019.

The Environment Dimension

The environment dimension describes the surroundings in which the bot operates. This refers to the machine or server that the bot inhabits, the network(s) that the bot may access, and the data systems that the bot manipulates. The environment category designates the bot's ability to access other systems and determines whether certain risks will emerge. For instance, parameters controlling type, scope, and closure control where the bot is located (such as a stand-alone machine or platform), how large the environment is, and any restrictions placed on the bot's access privileges. Failure to properly constrain a bot's access to systems presents an enormous risk to security and the integrity of data within the bot's reach. For example, an unconstrained bot may manipulate or delete data that it should not.

Furthermore, the environment category includes descriptions of the other entities that could be acting within the same space and how a bot's actions might change other users' or bots' experiences. When developing a bot, USAF must consider who will be occupying the same space as the bot. Otherwise, a bot could lock out another user by leveraging a document or dataset while another user wants access. Alternatively, two bots manipulating the same dataset at the same time could initiate a feedback loop in which the bots will eternally respond to each other's actions, to the detriment of the processes they are performing.

The Intrinsic Dimension

The intrinsic category describes functional properties of the bot itself, where the bot's developer determines each attribute. This group of parameters comprises the bot's actions, what a bot knows or remembers, where it gets its information, how the information influences its behavior, and whether or how it learns from the information. Characteristics of the bot's knowledge, such as where information comes from and whether the bot can remember past events, determine how the bot will behave. For instance, a bot's knowledge could be directly encoded within its programming or

retrieved from an external file. If it comes from an external file, it is vital to monitor who has access to the file and to document when and how the file changes. Otherwise, the bot might behave in an unexpected manner and users will either notice and be unable to explain results or, worse, not notice and have confidence in faulty results.

The intrinsic category also frames the outcome that a bot is working toward. One intrinsic outcome is goal orientation, which is characterized by many attributes, including complexity, attainability, and delegation, where *delegation* refers to whether the bot has authority to act on behalf of or even pretend to be the user. Current RPA bots within USAF act as though they are the user, leveraging the user's access credentials and CAC. This makes it impossible for a distinction to be made between the actions of the user and the bot. In cases in which a bot behaves unexpectedly and damages data or systems, it will be difficult to identify whether it was the user or bot causing the damage. The RPA Center of Excellence (COE) has requested funds for a centralized cloud that would track user and bot actions separately by giving the bot separate access credentials from those of the user.

The intrinsic category also includes reasoning, which determines how a bot responds to stimuli, the visibility of its actions, and whether a bot's actions require permission from a user. This collection of parameters is crucial to consider before bot development because they control the extent to which users can observe a bot's behavior and actively prevent a bot from performing certain actions. Considering standard guidance regarding actions a bot is able to take, how it documents its actions, and which bot actions require user approval will be beneficial.

Adaptability controls how a bot learns from its environment, restricting how the bot may adapt its behavior, the source that controls how a bot adapts, and whether there is any guidance given the bot's adaptation. This category is only relevant to IA bots that leverage AI or ML. There are no plans to develop IA bots within the A4 community.

Finally, the intrinsic category also includes anthropomorphic qualities in case the bot is given a visual representation or personality. For example, conversational bots that leverage AI to interact with users can be given a name, age, or ethnicity to personify them. This generally improves the user experience by allowing the bot to emulate human interactions. These types of bots are typically deployed for IT support and customer service functions.

The Interaction Dimension

The third and final dimension is interaction, which describes the rules governing how a bot engages with different entities in its environment. Specifically, the interaction category dictates the manner in which a user communicates with others. This includes how the bot inputs or outputs information and the language in which it communicates. Furthermore, the interaction category includes the ability of the bot to prevent and correct errors that it identifies in its inputs. This is critically important to bot development. After all, a bot's inputs could be prone to human error. If a bot blindly acts on the erroneous input, there could be significant consequences. For example, suppose a user leverages a bot to order parts but accidentally orders 400 units instead of 40. A bot with error prevention capabilities would notice that the order greatly exceeds what is typical and would ask the user if they are sure they want that many.

Using this taxonomy, USAF can more clearly define the needed dimensions of bots prior to their development. Furthermore, USAF should use the taxonomy to set standard guidance on bot characteristics and the environments that they can inhabit. Without applying a taxonomy, USAF could run the risk of relying on responsive rulemaking that identifies bot issues only after a problem occurs. This is of particular concern for the status of bot development in USAF because of the decentralized nature of bot development. The status of bot development in USAF and in commercial industry is discussed in more detail in the next section.

Current Status of Bot Development

USAF airmen have been burdened by tasks that are time intensive, are repetitive, and require high levels of consistency. In response to these challenges, USAF has considered opportunities to introduce bots to USAF processes and free airmen's time for other tasks. Faced with similar challenges, commercial industries have developed bot technologies to improve efficiency and increase consistency with varying levels of success. In this section, we discuss the initial steps taken by USAF to introduce bots, including the creation of the RPA COE and the training of *citizen airmen developers*, a term used by USAF to describe its targeted community of bot developers.[99] We also present the status of bot development as a service, including companies with which USAF has some existing partnerships. Finally, we describe the experiences of bot development in supply chain and logistics companies. These companies, although largely in the early stages of bot development, provide a road map of possible avenues for future bot development and critical considerations for bot investments.

Current Status of Bots in USAF

RPA bots are currently being developed in decentralized pockets of USAF. As a result, an effort is underway to centralize bot development and storage across the enterprise. At the core of this centralization effort is the creation of the RPA COE, which builds bots for stakeholders across USAF. The RPA COE has embarked on a partnership with UiPath, an RPA service provider and leader in bot development. This partnership includes the purchase of 50,000 licenses that allow users to develop their own bots for custom tasks with limited coding knowledge required. UiPath and other companies that offer RPA bot development services are described in more detail later in this chapter.

The development of the COE is well underway and has been focused on training airmen throughout USAF to become citizen airmen developers or bot developers who create simple bot applications for their personal use. Since its inception, the COE has trained more than 1,500 citizen airmen developers across 22 USAF functional areas at 90 installations.[100] The primary vehicle for this

[99] USAF's bot development outreach initiative is predicated on the notion that the best ideas for the application of bots will come from airmen across USAF (citizen airmen developers), who perform the repetitive, labor-intensive tasks that USAF seeks to automate through the use of bots.

[100] RPA Center of Excellence, "RPA Center of Excellence (COE) Accomplishments," PowerPoint presentation, March 2022.

expansion was a series of eight RPA roadshows and four RPA Digital Wingman Challenges.[101] These challenges, developed in partnership with UiPath, have trained more than 1,000 citizen airmen developers by providing initial bot development training followed by a competition in which airmen submit bot development ideas to a panel of judges.[102] Citizen airmen developers in USAF have designed 140 automations across 13 functional areas, with more than 50 bots in various stages of development.[103] Across USAF, the Secretary of the Air Force Financial Management & Comptroller (SAF/FM); Headquarters U.S. Air Force Deputy Chief of Staff for Manpower, Personnel and Services (HAF/A1); and HAF/A4 have led USAF with the highest number of developed automations by citizen airmen developers.[104]

To facilitate its citizen airmen developers, the COE is using its partnership with UiPath with the goal of making bot development tools available for every airman. UiPath RPA software is available through cloud access using CloudOne or by installing the desktop application.[105] CloudOne, the cloud computing system used by USAF to provide IT capabilities, hosts UiPath RPA software in an effort to centralize management and enable enterprise-wide access.[106] Furthermore, UiPath desktop software is available for installation to allow for bot development on local machines. Currently, around 50 RPA bots have been developed using UiPath, with varying success.[107] These bots are primarily deployed for personnel and finance applications.

Despite the COE's progress in partnering with UiPath and training citizen airmen developers, our discussions with USAF experts revealed that airmen have struggled to develop bots without major assistance from professional developers.[108] The COE also reports that a lack of funding has made scale-up efforts difficult, but, as of 2022, COE had submitted a request for funding to ensure that the center is capable of developing, scaling, and sustaining automations across USAF.[109]

Meanwhile, the A4 community is endeavoring to establish an organic RPA capability in the SCMW. This effort was separate from the RPA COE, but HAF/A4 has brought RPA COE "into the fold."[110] Although this program was still in its infancy at the time of writing, HAF/A4 was working toward developing bots to automate five specific processes, described in Table 3.6. SAF/FM uses a bot development life cycle framework that could be highly valuable for HAF/A4 to leverage when developing initial use cases.[111] This framework proceeds in the following seven stages:

---

[101] RPA Center of Excellence, 2022.

[102] Kayla Prather, "AFMC Robotic Process Automation Roadshow Drives Innovation," Air Force Materiel Command, December 15, 2021.

[103] RPA Center of Excellence, 2022.

[104] The list of bots we consulted was provided by RPA COE.

[105] RPA COE personnel, discussion with the authors, March 2022.

[106] Cloud One, homepage, U.S. Department of Defense, undated.

[107] The list of bots we consulted was provided by RPA COE.

[108] USAF 448th SCMW/Individual Mobilization Assistant, guided discussions with authors, May 2022.

[109] USAF AFLCMC/Business Enterprise Systems Directorate (GB)-Command, Control, Communications, Intelligence and Networks Directorate (HN), guided discussions with authors, March 2022.

[110] USAF Air Force Reserve Command 927th Wing, guided discussions with authors, July 2022.

[111] Air Force Financial Systems Operations, *AFFSO Automation Development Lifecycle*, PowerPoint presentation, February 12, 2020, Not available to the general public.

- intake: identify potential use cases
- prioritize: determine order of implementation for use cases according to strategic priorities
- design: capture the current-state process and design the automated target state
- develop: develop the bot based on the workflow design and requirements
- test: test to ensure consistency in expected outputs based on the requirements
- release: migrate code from development environments to a production state
- maintain: run, evaluate, and modify implemented automations to mitigate any risks that emerge during operation.

Table 3.6. Ongoing HAF/Logistics Directorate Bot Development Efforts

Process	Current Activities	RPA Use Case
CEMS—IMDS/G081 reconciliation	Manually reconcile APSR and MIS to validate the current engine status	Automate APSR and MIS reconciliations on a more routine frequency to identify variances requiring remediation
Stock control failed status	Manually reconcile BOSS alerted requisitions via their respective SoS to validate requisition and update order status	Automate BOSS alerted requisition reconciliation between ILS-S and SoS to validate requisition and update order and shipping status
437th SCOS ILS-S reports generation	Enterprise-wide process used by base users to request the scheduling of reports through 437th SCOS	Automate process to receive, validate, verify, and schedule user reports
EMB download process	Manually download, convert, and upload engine diagnostic data into maintenance systems and CEMS	Automate multiple data formatting steps and data upload process logging electronic records
435th SCOS/437th SCOS MRSP transfer	Cross-organizational process facilitating movement of accountable MRSP records	Automate process to receive, validate, verify, and execute movement of records

SOURCE: Features information from HAF/A4 Logistics Automation, 2021.
NOTE: CEMS = comprehensive engine management system; IMDS = integrated maintenance data system; G081 = a numbered government system for maintenance information for USAF airlift and aerial refueling weapon systems; APSR = accountable property system of record; MIS = management information system; BOSS = base operating stock specialist; ILS-S = integrated logistics system supply; SoS = source of supply; SCOS = supply chain operations squadron; EMB = engine management board; MRSP = mobility readiness spares package.

Even with the help of expert developers, bot development in USAF faces significant challenges, including incompatible systems and inefficient processes. Interviews with USAF experts revealed that at the supply wing alone, there are at least "three different platforms that do not communicate with each other," which translates to difficulties in centrally processing and sharing data coming from these platforms. Furthermore, replacing inefficient processes with RPA can magnify bottlenecks rather than resolve their underlying issues. These shortfalls are discussed in more detail later in this chapter.

Although RPA development in USAF is still in its infancy, IA bots have received even less attention. In our limited research, we became familiar with one IA bot prototyped by the Air Force Research Laboratory (AFRL), perhaps unsurprisingly because AFRL is mandated to be at the cutting edge of USAF technology development. AFRL prototyped a conversational bot through a contract opportunity called Multi-Source Exploitation Assistant for the Digital Enterprise.[112] This is the only actual USAF IA bot that we learned about, though there are likely others that have been tested. The objective of this AFRL project was to leverage AI/ML to create a bot that converses with the user as an assistant, answering questions or guiding the user to resources. An example of what the prototype bot can do is autonomously answer questions about commercial air traffic.

The overarching concept of IA bots holds significant promise for many applications beyond the scope of this particular AFRL effort. For instance, an IA bot could provide pilots with real-time plane diagnostics through a conversational assistant. Another application could be enabling maintenance personnel to ask questions about the correct type of parts to use and their availability. In general, IA bots facilitate improved content management where they can easily be leveraged to query reference material faster than a human. The extent of the promise of IA bots to USAF is likely very large, but will only be known through real-world projects, testing, and retrospective analyses. Improvement of data quality and migration of data to USAF's Basing and Logistics Analytics Data Environment (BLADE) might introduce new opportunities to leverage IA bots to inform strategic supply chain management decisions.

Although USAF is in earlier stages of RPA and IA bot development, commercial industry has developed these automations for diverse applications, from chatbots for customer service interactions to web crawlers that are the basis of search engines used to navigate the internet. In the next section, we discuss bots developed in two portions of commercial industry—RPAaaS and bot development by supply chains and logistics companies—and identify lessons from these experiences applied to USAF.

Informing USAF Bot Development with Experiences from Commercial Industry

Commercial industries were some of the earliest adopters of software bots to reduce personnel hours and achieve tasks beyond human capabilities. These software bots accomplish a wide variety of tasks; web crawlers serve as the backbone of online search engines and conversational AI, which provides services for millions in the form of Siri, Alexa, and Google Assistant. As USAF develops its software bots, commercial enterprise can provide useful lessons for valuation of bots, the application of software bots, and the future of bot capabilities.

In this section, we focus on two types of commercial bots: bots created and sold as a service (such as RPAaaS) and bots created by firms for internal use. Although bots are used for a variety of business aims, we focus our attention on supply chain and logistics commercial bot applications to provide insight on the A4 community's opportunities for applying bots to prevent supply chain

[112] "Multi-Source Exploitation Assistant for the Digital Enterprise (MEADE)," webpage, SAM.gov, last updated September 30, 2022.

degradation. From these commercial bots, we identify opportunities and challenges for implementing software bots and draw lessons for USAF.

Bot Value Proposition in Commercial Industry

With increased focus on efficiency and accuracy in complex global supply chains, companies have turned to software bots to reduce time spent on tedious and repetitive tasks while increasing the accuracy with which they are executed. RPA can automate these tasks without requiring a system change by repeating manual tasks at a higher speed. Without changing the system, RPA bots can continue to follow the rules and compliance created by the legacy IT system while reducing the cost to the organization by removing manual tasks that reduce productivity.[113]

These benefits are not without challenges. If a process is inherently flawed, accelerating that process could have the opposite effect by increasing the magnitude of errors. These shortfalls and their applicability to USAF are described in more detail later in this chapter.

RPAaaS

As described earlier in this chapter, the COE has contracted with UiPath, an RPAaaS provider that provides low-to-no-code environments that can be accessed with limited coding knowledge. Gartner, an IT consulting firm, identifies UiPath as a leader in RPAaaS for its strong "ability to execute" and "completeness of vision."[114] Gartner is a for-profit company that conducts technology market research. As part of this research, Gartner produces proprietary rankings of RPAaaS companies in its Magic Quadrant report, which ranks companies on two dimensions: ability to execute and completeness of vision.

The Gartner evaluation should be considered with some caution. Although the Gartner market research report was one of the only evaluations we identified, there is limited transparency for its methodology. For example, the publicly available methodology documents accessed through its website (without submitting a request or logging in) define terms used in the Magic Quadrant but do not include information on data collection, quantitative evaluation, or weighting that can be used to determine its objectivity.[115] Therefore, we present the Magic Quadrant for consideration of dimensions for evaluating RPAaaS and commonly highlighted RPA features but do not present Gartner's findings as a definitive ranking of these companies.

In its third annual Magic Quadrant report, Gartner classified and ranked 18 RPAaaS providers across the two dimensions. In addition to UiPath, the upper right quadrant of Figure 3.2 (maximizing both dimensions) includes three more top companies: Automation Anywhere, Blue Prism, and Microsoft.[116]

---

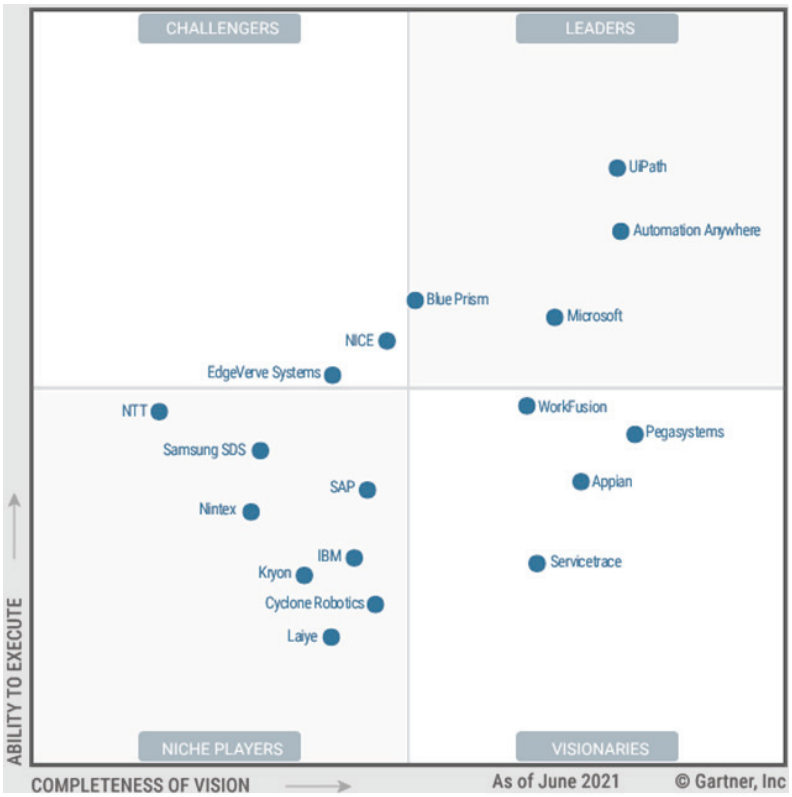
[113] Swapnil Sirdeshmukh, Yashdeep Saran, and Ankit Tondon, "Faster Decision-Making with RPA in High-Tech Supply Chains," Infosys, February 2, 2019.

[114] "2022 Gartner Magic Quadrant for Robotic Process Automation," webpage, UiPath, undated.

[115] We reviewed Gartner's methodology in documents on its webpage without using a login. See "Positioning Technology Players Within a Specific Market," webpage, Gartner, undated; Gartner, "Gartner Magic Quadrant and Critical Capabilities: Methodologies Evolution," September 10, 2019; and "Ombuds: Guiding Principles," webpage, Gartner, undated. Further supporting materials might be accessible through Gartner's portal, but this information might be proprietary.

[116] "2022 Gartner Magic Quadrant for Robotic Process Automation," undated.

Figure 3.2. Gartner's 2021 Magic Quadrant



SOURCE: Reproduced from "2022 Gartner Magic Quadrant for Robotic Process Automation," undated.
NOTE: Gartner surveyed business technologists to categorize companies across two dimensions: "ability to execute" and "completeness of vision." For more information, see "Positioning Technology Players Within a Specific Market," undated. These data have since been updated.

Gartner classifies RPAaaS leaders as companies that have reliable performance with a strong consumer base that enables them to lead the RPAaaS market toward innovation.[117] Each of the four leaders has some development of a low-code environment, which allows users with little coding experience to access a user interface to create the bot. The four leaders consistently offer automation for digital processes, and some allow task capture and mining, as well as ML and analytics. Leaders in RPAaaS are developing AI and ML tools to enhance repetitive processes with real-time decisionmaking. Although these tools might be useful for the A4 community's adoption, they require high-fidelity data to produce meaningful and accurate results. RPAaaS companies also plan to expand low- and no-code environments to make bot development more accessible to those without programming experience. Rather than using programming languages to have the bot conduct a specific task, the end user can use prebuilt templates to create a task flow.[118] A key takeaway for the

[117] "2022 Gartner Magic Quadrant for Robotic Process Automation," undated.

[118] Abhishek Shanbhag, "Why the Future of Chatbots Is Low Code," *BotCore by Acuvate* blog, January 15, 2021.

A4 community is that low- and no-code environments are an opportunity to use bot development more widely despite the limited programming expertise of airmen.[119]

Finally, bot development companies have worked to improve optical character recognition technologies. Optical character recognition will allow bots to take paper inputs (i.e., forms or printed materials) and incorporate them into process flows. In the current state, a person might need to input information manually from a physical source and incorporate it into digital resources. This process is prevalent in USAF because of the use of older data systems that are often disconnected from each other. Incorporating optical character recognition could allow the A4 community to bridge gaps between systems and reduce time spent on data entry.

Bots in the Commercial Supply Chain

Companies that provide RPAaaS often develop bot applications for customers with limited technology capabilities. Beyond these service providers, bots are being developed by logistics and supply chain companies for their internal use. Although these companies have strong incentives to implement automation, they often face challenges similar to those faced by USAF.

The rate of bot adoption by the logistics sector has been slower than in such other sectors as mining and warehousing.[120] In the logistics sector, container ports could have issues materializing the benefits of automation, including limited returns for operating expenses or failing to meet increases in productivity within expected time frames.[121] These challenges are due, in part, to issues filling specialized positions necessary to create and manage automation at ports and the additional training time required.[122] Furthermore, ports do not often have high-quality data that are structured in a transparent data pool that would enable sophisticated automation and AI technology.[123] Despite these challenges, *Supply Chain Quarterly* reported that it expected ports to pursue automation and AI because of the benefits to consistency, productivity, compliance, and safety.[124]

Although RPAaaS providers develop software bots for external use, some commercial supply chain and logistics companies develop software bots to resolve internal issues with production and workflow. These firms cope with similar supply chain challenges to USAF, including aging systems and the need for reliable decisionmaking. For example, like USAF, the commercial transportation and logistics industry uses manual documentation for supply chain management. Manually extracting data is estimated to result in an extraction error rate of up to 50 percent.[125] As outlined earlier in this section, the valuation of bot development should include the estimated savings from the expected reduction in errors.

---

[119] USAF AFLCMC/GB-HN, guided discussions with authors, March 2022.

[120] Fox Chu, Sven Gailus, Lisa Liu, and Liumin Ni, "The Future of Automated Ports," McKinsey & Company, December 4, 2018.

[121] Chu et al., 2018.

[122] Chu et al., 2018.

[123] Chu et al., 2018.

[124] "Maritime Port Operators See Great Promise in Artificial Intelligence," *Supply Chain Quarterly*, September 20, 2019.

[125] Gružauskas and Ragavan, 2020.

Using RPAaaS companies for bot development has allowed USAF to begin developing them without requiring a significant technology-trained workforce. However, using the bot license framework limits bot capabilities to individual initiatives, which might result in decentralized development by citizen airmen. This could limit USAF to creating bots for relatively simple tasks based on the problem sets faced by the airmen who obtain licenses. Orchestrating bot development or investing in bot development capabilities could enable the creation of bots that affect larger-scale business processes.

Using RPAaaS for bot development could also result in service lock whereby USAF needs to continue paying vendors to maintain and manage the automation. Decentralized bot development in USAF has resulted in the use of multiple automation platforms that do not communicate or work together.[126] USAF will need to settle on an approach: either use RPAaaS bots as a temporary measure and invest in its own bot capability or create business process changes that remove the need for the bot in the long term.

Supply chain and logistics companies that implemented their own bot development capability can face delays on ROI. The evidence suggests that these challenges were caused by limited data quality and understaffing of specialized technical positions.[127] We discuss these issues and their applicability for USAF in the next section of this chapter.

Limitations and Potential Risks for Bot Development in USAF

Bot development could allow USAF to make processes more efficient and free airmen's time for other tasks. But these benefits are not without challenges. In this section, we highlight the general limitations of RPA and situate these limitations within USAF operations. Some limitations to RPA need to be considered before developing a bot. First, RPA bots, as defined earlier in this chapter, require well-structured inputs and cannot read inputs that come in nonstandardized formats, handwritten documents, graphs, charts, or images.[128] Second, a lack of cognitive capability via AI prohibits analysis of trends or judgment calls to be made by the bot. Although IA overcomes some of these limitations, it also introduces new risks. We discuss these risks and other possible limitations of IA bots later in this section.

As discussed previously, RPA is often used to expedite processes within a broader system. But, in doing so, RPA can introduce new bottlenecks or magnify inefficiencies of a poorly designed system. Although not addressed in the bot taxonomy, if a system consists of several sequential processes, expediting one process could simply create an unwieldy backlog at the next step. Principles of continuous improvement tell us that a stable system can function at no greater speed than the speed of its bottlenecked process. Expediting any process before the bottleneck will only increase the backlog where the process is congested, and expediting processes after the bottleneck will only create

[126] RPA COE staff, guided discussions with authors, March 2022. Licenses currently in use by USAF include UiPath, Automation Anywhere, and BluePrism.

[127] Chu et al., 2018.

[128] Abhimanyu V, "What Are the Limitations of RPA?" Tutorialspoint, December 8, 2022.

excess unused capacity. The system will only improve if it is the bottleneck process itself that is expedited, potentially through automation.

Deploying RPA bots without proper testing incurs the risk of making errors much faster and on a larger scale than humanly possible, irreparably modifying or deleting large amounts of critical data. This risk is seen in a particular combination of features of the taxonomy—specifically, the predictability, agency, and visibility features of the intrinsic category that control the freedom given to a bot and the ability for others to track a bot's actions. Within the interaction category, the access and act features determine the bot's ability to make changes to specific systems, and robustness features control the bot's ability to prevent and correct errors that are detected in inputs before acting.

An organization could also suffer when the number of bots in use becomes unwieldy.[129] It can become difficult to pinpoint which actions were performed by which bot and predict such outcomes of bot interaction as when multiple bots manipulate the same data. This potential risk should be identified in the population and diversity criteria in the environment category of the taxonomy.

Bots increase the complexity of any system through additional software requirements, where different bot development software carries technology specifications that might not be compatible with certain data systems or work with other bot software. It is a good practice to choose one bot software that works on all critical systems for the full enterprise. The type of software used to develop the bots is captured in the type, scope, and closure features of the environment category in the bot taxonomy.

As some airmen duties are replaced by bots, it is important to develop plans defining changes in the scope of jobs, how time will be repurposed to other tasks, how management might change, and how metrics of job performance might change. Some individuals might fear that they are being replaced by a bot. Without buy-in, HAF and even some airmen might resist an automation initiative.[130] Furthermore, any changes to processes need to be approved by all process owners and supervisors so that everyone is aware of changes to processes and the broader system.

In addition to the process risks mentioned above, bots present significant cybersecurity risk when not implemented or managed carefully. Bots are given system access credentials and the ability to move, modify, and delete data at high speeds. Malicious or negligent actions can also be difficult to track because it is often difficult to separate the actions of a bot from those of a user.

> Poorly managed bots can negatively impact the confidentiality, integrity or availability of the information stored and processed by an organization. This applies not only to the infrastructure components that support the RPA environment—servers, databases, virtual machines, and orchestration technology—but also to the passwords and permissions of the accounts that bots use to interact with applications and systems.[131]

Chapter 4 discusses cybersecurity issues related to integrity-based attacks with special emphasis on bots. It presents an approach for identifying and mitigating these types of cybersecurity risks.

---

[129] "Maritime Port Operators See Great Promise in Artificial Intelligence," 2019.

[130] Tony Abel and Ben Franjesevic, "Who Is Watching the Bots? Part 2: Operational Challenges (and Solutions)," Protiviti, June 11, 2019.

[131] Abel and Franjesevic, 2019.

Building and Deploying Robotic Process Automation Within USAF

Because RPA development efforts in USAF have been largely stovepiped, there have been numerous instances of rogue operators building and deploying bots without getting approval to do so and without having proper bot oversight. This poses an enormous risk to the integrity and security of critical systems. Because these bots live on local machines, supervisors often cannot differentiate the actions of a bot from those of the user. Indicators of bot actions include different rates and timing of transactions (bots can run continuously and all day). To centralize management of bot operations, the RPA COE is working toward hosting bots centrally on CloudOne.[132] This repository will enable better collaboration among developers and better access to airmen who wish to leverage existing automations. There are many advantages to a centralized approach. For example, to better differentiate bot and human activity within a user's account, the COE wants to add distinct human and bot tokens to CloudOne to monitor bots and human actions separately. (The COE is waiting on $400,000 in funding to implement this modification.)

As mentioned earlier, RPA development requires specialized technical expertise that is not developed within USAF. For example, during interviews, we were told that, "Not a single bot was built by an airman that could function. From looking at over 100 bots, the only bots that have made it to production were rebuilt by professionals."[133] In a news release for the 2019 Air Force Science and Technology Strategy, the science and engineering division highlighted that fill rates for USAF positions requiring advanced science, technology, engineering, and mathematics (STEM) degrees were low, often around 50 percent.[134] A common suggestion from experts in USAF was that USAF should focus on recruiting airmen with more knowledge of technology and mathematics.[135] They described the ideal prerequisite training for future bot developers as needing to include operations research, data science, computer science, and electrical engineering: "Another goal is to build the skills of the airmen for the 21st century. We have massive data systems but not the analyst culture to use it."[136]

Finally, there are concerns regarding the ability to scale up bot development and deployment, as no standardized guidance exists on how to leverage massively scaled services.

Proposed Bot Applications for the USAF A4 Community

In the preceding sections of this chapter, we presented ways to define a bot, ways to characterize a bot, and general use cases for both RPA and IA bots. In this section, we suggest four categories of

[132] USAF AFLCMC/GB-HN, guided discussions with authors, March 2022.

[133] USAF AFLCMC/GB-HN, guided discussions with authors, March 2022.

[134] Amanda Miller, "Half of Air Force Advanced STEM Billets Go Unfilled or Require Waivers," *Air Force Magazine*, August 21, 2022.

[135] RAND SMEs, guided discussions with authors, August 2022; Miller, 2022.

[136] USAF AFLCMC/GB-HN, guided discussions with authors, March 2022.

bots, listed below, that might have use for the A4 community. We deductively developed the following four categories, synthesizing what we learned from our research:

- bots that expedite high-volume processes
- bots that execute new processes
- bots that link disparate data systems
- bots that facilitate robust analyses.

We also suggest a variety of possible applications that we believe would be useful to the A4 community that correspond to these four categories. We compiled these opportunities for bots through guided discussions with RAND and USAF SMEs. These potential bots are listed in Table 3.7. We then describe a demand forecasting–related bot as an illustrative use case.

Table 3.7. Potential Robotic Process Automation Applications for USAF Implementation

Application	Description
Expedite high-volume processes	
Submit work orders	Automatically submit work orders using a bot linked to an Electronic Technical Manual (ETM). The user selects a part from an ETM and the system automatically submits requisition and updates statuses on maintenance systems.
Generate regular briefings for different levels of command	One bot could use one set of information to generate multiple briefings for different levels of command by varying the resolution of information presented. For example, one tactical briefing for maintainers might present part availability information. A separate briefing for commanders could be a strategic view of systemic issues. Both briefings could be constructed by a bot using the same data.
Generate management reports	A bot could create management reports that show segmentation of mission impaired capability awaiting parts (MICAP) response times, including such segments as (1) MICAP verification, (2) sourcing time, (3) Pull/Pack/Ship, (4) transportation time, (5) base processing time, and (6) receipt by maintenance.
Execute new processes	
Maintain comprehensive status updates	Report open work orders, status on parts, and equipment status reports throughout supply chain. Requires getting data from other information systems in DLA to find status of requisition outside USAF.
Build empirical bill of materials	Create reference material by linking maintenance actions (by serial number, work orders, etc.) to understand what parts are needed to sustain and maintain equipment. A bot could also identify superfluous parts in the system that do not belong in the bill of materials for any weapon system.
Track the demand of parts over time	Dynamically monitor part failure rates. Identify parts where failure rates change suddenly. Link back to vender or maintenance systems to identify root causes.

Application	Description
Evaluate the health of the supply chain	Detect where queues are building on arcs or nodes within distribution pipeline. Gather information on status of shipments to figure out which nodes or arcs are causing problems.
Identify nontraditional suppliers	When a program office is going through a nontraditional supplier, a bot might help evaluate the best available alternatives (e.g. using eBay because a part is not available from the traditional supplier).
Catch errors	A bot could monitor a user's inputs for potential mistakes and prompt the user to correct them (e.g., "Are you sure you meant to order this part" or "you only entered 7 digits; please correct").
Ensure data integrity	Bots that can cross-reference other systems to perform a data integrity check (e.g., MICAP requisitions will show closed in D035A but open in ILS-S and accrue excessive MICAP hours).
Link disparate data systems	
Clean and consolidate data within a data lake	Bots can perform cleansing on data flowing to the data lake so that everyone can use the same clean data source. All information systems send data into data lake with development environment (includes Python, R, structured query language) to link data across systems (R is a programming language commonly used for statistical analysis and data visualization).
Connect DLA and USAF supply systems	Examine DLA's and USAF's systems for all of the low-demand, high-cost parts to document inventory levels. See how many planes are waiting on a single part.
Facilitate robust analyses	
Collect additional data to assess supply chain risk	Scrape information provided by suppliers about specific components received for each aircraft. Compile worksheet to link part to part and connect to public finance datasets to evaluate risks. Have information on all parts coming in from various sources.

SOURCE: Features information from RAND SMEs, discussions with authors, August 2022.
NOTE: These processes are described as possible bot applications but should be fully evaluated to ensure that the potential benefits outweigh introduced risks. We discuss these trade-offs in more detail in Chapter 4.

Proposed Use Case

As discussed in Chapter 2, a prerequisite for improved causal demand forecasting is better integration of part demand, usage, and maintenance data. We chose this example for several reasons. First, the datasets are not linked, making it difficult for USAF personnel to get a holistic view of parts failures, actions that occur within the supply functional community because of a part failure, and actions taken within the maintenance functional community that could provide additional insights about the cause of the failure. Second, there are pockets of personnel within USAF that have manually linked and cleansed the data for the purposes of doing the type of analysis we mention.

Third, the concept for the bot is not one that the A4 community might think of given its functionally stovepiped approach.

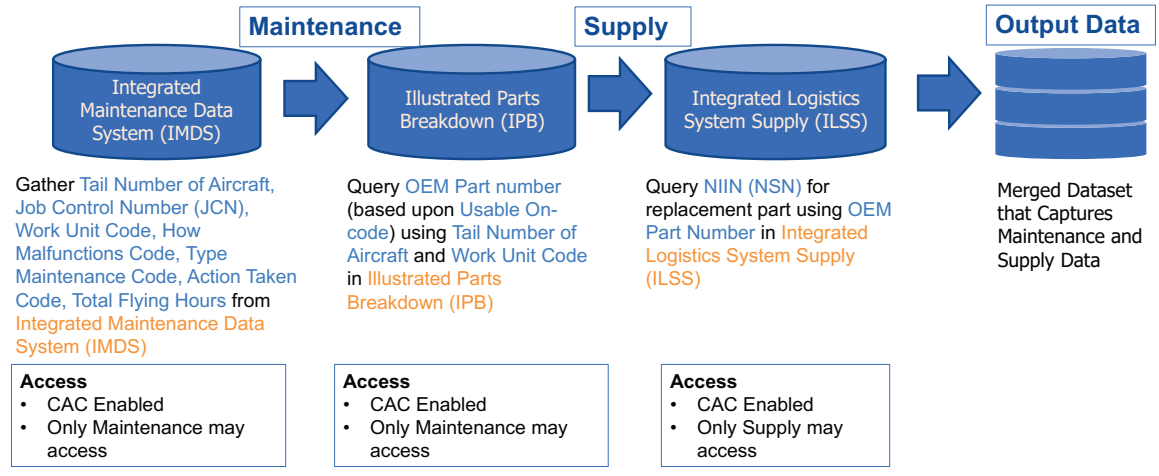
Although research has not conclusively shown that linking these datasets can yield better forecasting, such integrated data are required to investigate potential benefits. In addition to demand forecasting, linking supply and maintenance data might yield other benefits that will be discussed later. Given the potential benefits of linking supply and maintenance data, in this section we present a theoretical bot that can link data between the supply and maintenance systems.

The information resides on three individual systems: one for maintenance activities, one to correlate part numbers, and one for supply activities. There is no easy method to bridge these data sources. However, it is possible to design a bot to integrate the three datasets to enable more robust analysis to drive supply chain operations and inventory decisions (shown in Figure 3.3).

The bot would function in the following three primary steps:

1. The user gathers the following inputs from IMDS:
   - tail number of aircraft
   - job control number (JCN)
   - work unit code (WUC)
   - malfunction codes
   - type maintenance code
   - action taken code
   - total flying hours.
2. Leveraging the tail number of the aircraft and the WUC, the bot queries the original equipment manufacturer (OEM) part number from the illustrated parts breakdown (IPB).
3. Leveraging the OEM part number, the bot queries the national item identification number or the national stock number (NSN) information from ILS-S.

Figure 3.3. Proposed Use-Case Automation



NOTE: Total flying hours in IMDS are at the time of failure.

The output would be a merged dataset that captures maintenance and supply system data, which is placed in a data repository and could facilitate more robust analyses. The use case we propose could have a variety of benefits, elaborated on in the following sections. (Note that we do not explicitly demonstrate these benefits.[137]) First, associating the JCN with a WUC and an NSN could allow for a more nuanced analysis of what is driving aircraft downtime because it could account for additional variables. For specific parts and weapon systems, noting the calendar date and measuring flying hours between each replacement and the total flying hours at failure could provide data to validate distribution of failures and to understand implications of failures over time.

Second, it could better inform repair decisions. Repair decisions are currently made based on cost of repair at a depot, after pulling the part, testing it, and sending it off to arrive at a decision on the total cost of repair. However, this decision can only be made with data systems that link man-hour costs (IMDS) with the activities that led to requisition from the supply depot, the part being sent back, and the cost of repair at the depot. Furthermore, if the elapsed time between replacement is tracked for each new system at the tail number level, these early failures can be used to predict future performance more accurately than if engineering analysis is used to predict future performance. In new aircraft and modifications, it can be difficult to determine the total population of operating systems, which complicates traditional methods to calculate flying hours between failure. Measuring flying hours between replacement for each failure is one method to make early predictions for system performance.[138]

Third, understanding maintenance man-hours (MMH) required to first troubleshoot and then remove and replace a bad component could provide insights into training shortfalls or deficiencies. Moreover, component costs coupled with MMH costs and aircraft downtime could provide insights into ways to improve aircraft availability. Finally, identifying the lowest replacement level could reduce DLR replacements for major systems when subcomponents can be replaced. In Chapter 4, we apply a cyber risk framework to identify specific cyber risks.

USAF is conducting three related but distinct ongoing automation efforts. The first is a *Part Bot* that is being developed to procure parts.[139] The user inputs an OEM part number, and the Part Bot looks up the associated NSN and submits a requisition. This bot is in the beta testing phase before full deployment.

Although both the Part Bot and our proposed use case query an NSN number using an OEM part number, there are two differences between the Part Bot and the use case proposed here. First, the Part Bot is intended to order parts, while the use case is intended to facilitate robust analyses and a better understanding of linkages between maintenance and supply. Second, our use case adds the bridge from IMDS to IPB, automating the querying of the OEM part number, which the Part Bot

[137] Chapter 2 expands on the difficulty of doing this type of analysis today.

[138] For example, the Duane method uses a log/log plot of accumulated flying hours versus failures to predict flying hours between failures. Typically used in test and evaluation, it can also be used to predict in-service mean time between failures (J. T. Duane, "Learning Curve Approach to Reliability Monitoring," *Institute of Electrical and Electronic Engineers Transactions on Aerospace*, Vol. 2, No. 2, April 1964).

[139] Douglas Sangster and Benjamin Young, "PartBot (for Maintainers, by Maintainers)," PowerPoint presentation, 35th Maintenance Group, undated.

assumes the user will do manually. For our use case, the user inputs the tail number of the aircraft, the JCN, and the WUC, and the OEM part number is automatically queried.

The second related effort is ongoing at AFLCMC.[140] This is a mapping of WUC to NIIN. Although this bot does fully map sections of maintenance and supply systems, this work is limited to the F-16 and B-2 weapon systems, which have reference designators in the system that make them easier to map. Alternatively, the use case we propose will add an entirely new capability, linking JCN to part number to WUC to NSN to MMH. This will provide a more holistic view of a given repair by incorporating supply and maintenance perspectives.

Third, HAF/A4 has funded an automation that compares data from IMDS and the CEMS. The bot will identify discrepancies in location and engine status between the two systems and list these discrepancies in a report. This bot is intended to identify errors rather than bridge multiple data systems.

Applying the Taxonomy to the Use Case

Developing a bot to connect supply and maintenance systems would enable new analysis and provide a more nuanced view of repairs. As mentioned earlier in this chapter, the first step for bot development is to clearly define the bot's characteristics. To that end, we applied the taxonomy described above to characterize and define the use case we developed. This section is intended to be illustrative of the usefulness of applying the proposed taxonomy for USAF. This process allows for bot development that fully characterizes the bot before it is built.

Applying the taxonomy serves several purposes. The characterization

- allows for identification of possible risks, benefits, and limitations of the bot that might not be apparent from the bot's stated objectives as outlined in the previous section
- facilitates a proactive holistic perspective of the constraints placed on the bot in terms of systems that the bot can access and the tasks that the bot is allowed to execute, which in turn can reduce the risk of relying on responsive rulemaking that identifies bot issues only after a problem occurs
- informs a more thorough cyber risk analysis, which is discussed in more detail in the following chapter, where we use this specific bot example as one of our cyber case studies.

In the following three tables, we outline the characteristics of the proposed use case across the three dimensions described earlier in our introduction to the taxonomy. First, we begin with a description of the bot's environment, which consists of three data systems from which data are drawn and one system that outputs consolidated data. Next, we describe the bot's reasoning and the output data that represent the bot's goal. Finally, we describe how the bot will interact with the systems it accesses. In some cases, we draw a distinction between the bot's initial state (i.e., the first uses of the bot after it is developed) and finished state (i.e., the matured version of the bot that advances in sophistication).

[140] AFLCMC/GB-HN, guided discussion with authors, March 2022.

In Table 3.8, we apply the environment dimension to characterize the bot's ability to access other systems. This table will further inform the cyber risk analysis and can help determine whether certain risks will emerge. The reader will note that this closed environment includes IMDS, IPB, and ILS-S with the output placed on such a system as LIMS-EV, a testing environment, or a data lake. It is expected that the population will be both human users and bots.

Table 3.8. Application of the Environment Dimension Taxonomy to the Use Case

Characteristic	Potential Values	Description of Value
Type	Stand-alone	Initial state: could be on developer's local machine
	Platform	Initial state: could be in bot development sandbox Finished state: output data made available in LIMS-EV. Bot may be hosted on the Expeditionary Combat Support System (ECSS) at Gunter Air Force Base.
Scope	Bounded	Environment limited by capacities of system housing bot, IMDS, IPB, and ILS-S Note: Restrictions on what the bot can do. Might be unable to access sensitive data on these systems (e.g., strategic nuclear force).
Closure	Closed	The system housing bot, IMDS, IPB, and ILS-S (and LIMS-EV) are all CAC restricted and impose additional user restrictions on who can access
Dynamism	Dynamic	Users and bots may manipulate the system housing bot, IMDS, IPB, and ILS-S (and LIMS-EV)
Permanence	Episodic	The bot retrieves and consolidates data. Only change to environment is output of a log including consolidated data. However, no data systems are changed and no data are modified.
Population	Countable	Users of IMDS, IPB, and ILS-S (and LIMS-EV if output is declassified) are reasonably countable (i.e., the number of people who can access these systems). IMDS and ILS-S have especially large numbers of users. The large population makes it difficult to identify who made errors.
Diversity	Heterogeneous	The population could be made up of humans or bots.

The bot's intrinsic characteristics determine the functional properties of the bot itself. These characteristics for the proposed use case are outlined in Table 3.9. Once a user hits "go," the bot's knowledge consists of the data it pulls from IMDS, IPB, and ILS-S. Its reasoning follows simple if-then-else logic to query and match field values across systems, making it single tasked and predictable. Its goal is internal, fairly complex, and risky, where the data being compiled are large and important. The bot self-terminates when the data have been consolidated, and there is no anthropomorphism involved because the bot does not have a physical or audio representation.

Table 3.9. Application of the Intrinsic Dimension Taxonomy to the Use Case

Characteristic	Potential Values	Description of Value
Knowledge		
Memory	Short-term	It retrieves the data that are available, stores them temporarily, and then consolidates them based on identifying variables (primary keys).
Source	Supplied	Data supplied by IMDS, IPB, ILS-S
Reasoning		
Mechanisms	Scripted	Bot will be rule-based (if-then-else) with standardized inputs and predetermined outputs.
Agency	None	Initial state: require user approval before executing key steps
	Complete	Final state: bot runs autonomously, retrieving and consolidating data at predefined intervals in the background
Predictability	Deterministic	The same inputs will result in the same outputs.
Visibility	Transparent	The retrieval of data will be invisible, but the output log will be created and will show where the data came from.
Reactivity	Synchronous	No intentional delays will be implemented.
Scheduling	Single tasked	A user hits go, and the bot executes a single task.
	Multiple	
Adaptability		Not applicable: no learning behavior
Goal orientation		
Complexity	High	Compiles data from multiple sources that are not readily available
Criticality	High	Tasks are not high risk or critical to operations. However, relaxed data processing will synthesize a lot of important data in one place; could pose new security risks and might show vulnerabilities to supply chain.
Attainability	Achievable	Goals are well defined and feasibly reached.
Explicitness	Explicit	Output log with merged data is explicitly defined
Source	Internal	Goals are programmed into the bot.
Delegation	Complete	Depends on state of system where the bot is deployed. For example, it is impossible to distinguish bot from user on ECSS, IMDS, IPB, and ILS-S. However, a system owner may issue separate tokens for bot and user. The RPA COE is currently working toward hosting bots on Cloud One, which will distinguish bot from user.

Characteristic	Potential Values	Description of Value
Specialization	Specialist bot	Limited to specific portions of three data systems
Anthropomorphism		
Name	None	No name
Embodiment	None	No embodiment
Age	None	No age
Gender	None	No gender
Ethnicity	None	No ethnicity
Profession	None	No profession
Personality	None	No personality
Emotions	None	No emotions
Life cycle		
Lifespan	Terminating	The bot will self-terminate when the datasets have been merged and a log has been compiled
Creation	Human	The bot was created by a human
Reproduction	None	This bot does not create other bots

We apply the interaction dimension to the use case in Table 3.10. In this table, we outline the rules governing how a bot engages with different entities in its environment. The bot's access is partial because it is restricted only to certain fields and data systems. The bot is non-sensing but does act on its environment. Communication is limited to input and output of data. The bot would initially have no error prevention capabilities, but these could be added to prevent user error when entering inputs.

Table 3.10. Application of the Interaction Dimension Taxonomy to the Use Case

Characteristic	Potential Values	Description of Value
Access	Partial	The bot is restricted to only querying the fields (primary keys) necessary to bridge IMDS, IPB, and ILS-S.
Sense	Non-sensing	The bot cannot perceive environmental stimuli except when a user hits go. The bot has no physical hardware to interact with the environment.
Act	Acting	The bot does not manipulate data and does not interact with users but does retrieve and compile data to a single source.
Communicate		
Disposition	Indifferent	The bot is indifferent to others' actions.
Veracity	Truthful	The bot presents factual data without deception.

Characteristic	Potential Values	Description of Value
Cardinality	One-one	Even if multiple users can run the bot, only one user will run it at a time.
Directionality	Two-way	The bot inputs and outputs datasets.
Directness	Direct	The only communication would be an error message to the user, and the log is output directly by the bot.
Language capability	None	The bot has no language capability.
Initiative	Reactive	Waits until someone hits go.
Robustness		
Error prevention	User	Errors in the input data are the responsibility of users.
Error correction	User	The bot will provide error prompts but relies on users to fix errors in input data.
Mobility	Static	The bot retrieves data from other systems, but its code will always remain in the same place.

Observations and Findings

Bots and automation provide valuable tools for USAF to improve processes and limit supply chain degradation. In this section, we outline the following findings and observations from the bot analysis.

- The USAF A4 community is in its early engagement with bots—developing them in generally stovepiped functional areas to automate manual tasks. However, this approach does not allow the A4 community to fully leverage the potential of bots. For example, bots that tie together data from disparate elements of the logistics enterprise, such as those needed to provide a more holistic analysis of parts failures, could be highly impactful. Similarly, incorporating optical character recognition could allow the A4 community to bridge gaps between systems and reduce time spent on data entry.
- Early indications reveal that citizen airmen developers are challenged to produce usable bots. Questions remain about whether USAF personnel possess the technical expertise to fully leverage bot technology, and the data suggest that this concern is warranted. Low- and no-code environments are an opportunity to use bot development more widely despite the limited programming expertise of citizen airmen developers.
- Developing and implementing bots needs to be weighed against the cybersecurity vulnerabilities they introduce. Accordingly, the use of bots versus a modification to an information system to enable the capability that bots could provide needs to be weighed relative to the value the bots provide in both the short term and the long term.
- Unified direction and guidance regarding bots could help USAF in general and the A4 community in particular best maximize the potential of bots. Operating from standard policy that outlines what actions a bot is able to take, how it documents its actions, and which bot actions require user approval could be useful. Additionally, using a taxonomy to set standard

guidance on bot characteristics and the environments that they can inhabit could reduce the risk of relying on responsive rulemaking that identifies bot issues only after a problem occurs.

Chapter 4

Addressing Vulnerabilities of Cyber Tampering

As discussed in Chapter 3, the application of bots presents both opportunities and risks to the HAF/A4 mission. From a security engineering standpoint, this is true for any change in system design or composition: The introduction of new technologies must be considered through a risk-based lens to present decisionmakers with a balanced argument that considers both positive and negative effects. For example, a bot that accesses records to look for anomalies could present itself as an attractive automated security enhancement. However, such an addition must consider a host of challenges, such as adequately defining anomalies, separating valid changes from those presenting a threat, and guarding against the opportunity for the bot itself to serve as a vector for attack. Alternatives (such as, in this case, record-based encryption or signing) should then be examined in the context of technical, procedural, and training challenges that would accompany such a change. As applied to cybersecurity, these considerations are often referred to as requiring a mission assurance approach: not focusing cybersecurity decisions on individual systems or controls, but placing them in the context of overall mission risk. Doing so requires methods that consider the entirety of the system while supporting specific security actions.

The analysis to enable this level of cybersecurity decisionmaking forms the basis for the cyber tampering task of this effort. We were asked to focus on cyber threats that affect the integrity of the HAF/A4 mission, with emphasis on events that could easily go unnoticed (yet lead to major issues) and methods for mitigation.[141] It was essential that the process employed be equally applicable across the HAF/A4 enterprise, including legacy systems, systems undergoing modernization (such as the ESCAPE system), and future systems under consideration (such as those discussed in the prior chapter on bots). Most importantly, the analysis must place cyber risk in a mission context to inform decisionmakers of how to balance the risks and rewards for any system change and tie recommendations to the functions and data essential to HAF/A4 mission execution.

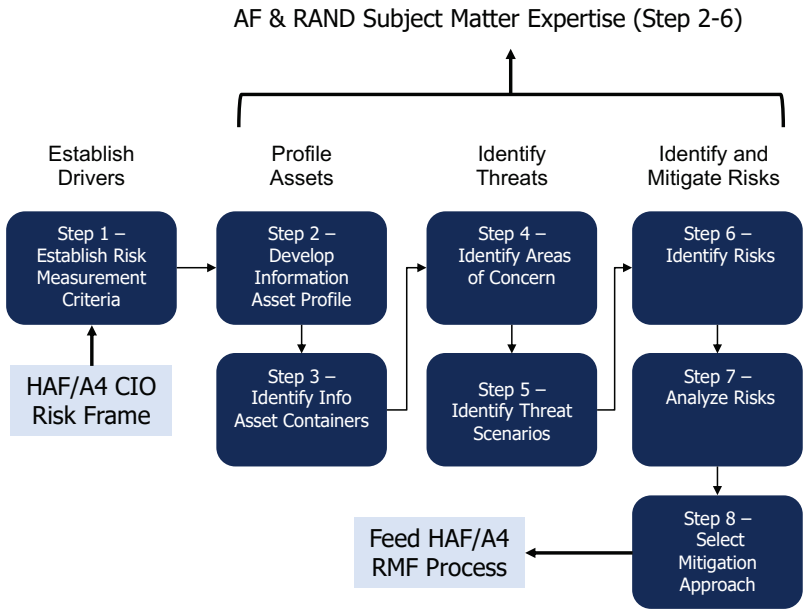
This chapter presents a process, based on established security engineering literature, to assist decisionmakers in systematically deciding where to invest in mitigation measures to reduce such risks as tampering. In addition, examples of the application of the process are provided to demonstrate how such a process can inform both specific actions (e.g., investment into specific controls) and general actions across the system life cycle.

[141] As previously discussed, this category of threats was informally referred to by the sponsor as *mischief versus mayhem*: cyber actors causing small, hard-to-detect problems that could have major downstream ramifications. Traditionally, cybersecurity considers risks relative to confidentiality, availability, and integrity, with tampering and errors a subset of the latter ("Tampering," webpage, National Institute of Standards and Technology Information Technology Laboratory Computer Security Resource Center, undated).

Methodology

For this analysis, we employed the OCTAVE Allegro process.[142] Developed by the Software Engineering Institute at Carnegie Mellon University, the OCTAVE methodology met our requirements as a structured, risk-based, full–life cycle approach to analyzing systems for the purpose of identifying potential security concerns (leading to the evaluation and selection of mitigations). The OCTAVE process unfolds as a series of eight steps, depicted in Figure 4.1.

Figure 4.1. The OCTAVE Allegro Process, with RAND and HAF/A4 Intersections Denoted



NOTE: CIO = chief information officer; RMF = risk management framework.

The core OCTAVE Allegro process steps proceed as follows:

- **Step 1. Establish Risk Management Criteria**: The first step in the process is the establishment of risk criteria. This step is purposefully disconnected from those that immediately follow to ensure that the criteria selected are independent of the system considerations to follow.
- **Step 2. Develop Information Asset Profile**: Central to the process is the act of identifying what are termed *information assets*. Information assets are key data elements necessary for the execution of the mission. Focusing on these assets allows emphasis to shift from the

[142] *Allegro* refers to a version of the process that is less process intensive to be streamlined and adaptable to a variety of uses. See Richard A. Caralli, James F. Stevens, Lisa R. Young, and William R. Wilson, *Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process*, Software Engineering Institute, Carnegie Mellon University, May 2007; and Christopher Alberts, Peter Gordon, and Audrey Dorofee, *Managing Information Security Risks: The OCTAVE (SM) Approach*, Addison-Wesley Professional, 2002.

instantiation of the specific architecture and placed on the information life cycle.[143] Doing so assists in setting the scope of the evaluation because only the containers (and associated threats) relevant to identified information assets are considered.

- **Step 3. Identify Information Asset Containers**: The third step of the process concerns the mapping of the more abstract information assets to the tangible realization of computing infrastructure. This involves enumerating the locations within the system where the information assets under examination reside, considering technical, physical, and people categorizations. This exercise ensures that the scope of the analysis includes both commonly considered elements (such as endpoints and network devices) and less common elements (such as backup devices, nonelectronic media, and key personnel). Where the examination is focused on existing systems (or systems under acquisition) for which documentation exists, these can be employed to support this mapping.
- **Steps 4 and 5. Identify Areas of Concern and Identify Threat Scenarios**: Steps 4 and 5 of the process identify potential areas of concern (Step 4), which are developed into system-specific threat scenarios (Step 5). As with other OCTAVE process steps, these may be performed using any number of widely accepted modeling and analysis practices from the security engineering community. In our experience, these steps can occur in a parallel, sequential, or iterative manner, depending on the analysis technique employed.
- **Steps 6 and 7. Identify Risks and Analyze Risks**: Step 6 establishes potential consequences for each threat scenario by considering its effects on information assets. Step 7 furthers this analysis by applying a quantitative measure for the risk assessment based on the risk metrics established in Step 1. The result is an impact (consequence) score for use in risk calculation, along with threat and vulnerability likelihood scores, to generate an overall risk score using the following definition of *risk*:

$$risk = [threat \times vulnerability] \times consequence.$$

  This separation of threat and vulnerability allows the analysis to consider both extrinsic factors (probabilities across actors or threat) and intrinsic factors (the presence of adverse conditions or vulnerability weaknesses, as identified in Steps 4 and 5). Although these steps lead to a quantitative measure, the output score denotes a relative importance of each risk. Therefore, the absolute numbers are less important to the result than their comparative values, which can be used to focus attention onto the greatest risks. As a result, OCTAVE Allegro suggests the practice of grouping risks together into risk pools, which often contain similar themes or concerns. Pools are tagged as having primary and secondary risk actions (mitigate, defer, accept, transfer), thereby guiding the final step.
- **Step 8. Select Mitigation Approach**: In Step 8, risks are examined for adjudication. Those that fall into pools that require action (i.e., mitigation) are examined for the addition of a

[143] The *information life cycle* is composed of information generation, processing, storage, communication, consumption, and destruction. More details on the role of this life cycle in the broader context of information assurance can be found in Kamal Jabbour and Sarah Muccio, "The Science of Mission Assurance," *Journal of Strategic Security*, Vol. 4, No. 2, 2011.

control or execution of a security engineering activity to bring the risk to a more acceptable level.

Ideally, each step of OCTAVE (and similar types of security analysis) is conducted as a workshop, bringing together process and domain experts to perform the analysis. For this research, we pulled from internal expertise in cybersecurity, risk management, and information from various real and projected Department of the Air Force (DAF) systems to perform and validate the analysis. Incorporation of a broader set of stakeholders in a workshop during future iterations would better refine the outcomes to HAF/A4 needs.

Further details regarding the advantages OCTAVE Allegro and similar processes provide and the rationale for our selection are presented in Appendix D. The choice of OCTAVE corresponded to our desire to employ a mission-focused approach compatible with existing processes (such as the DoD RMF)[144] that was applicable to legacy, modernizing, and emerging systems. Crucially, it supplied a repeatable, risk-based, life cycle-oriented process foundation on which we could develop our analysis.

Adaptation of OCTAVE to Support HAF/A4

Taking advantage of the structure OCTAVE Allegro (and similar processes) impose on the otherwise unstructured application of the various metrics, models, and best practices common in the security community, we adapted the process to better support security investment decisionmaking in a mission context. This section covers the approach taken to tailor OCTAVE Allegro to focus on the information integrity challenges of interest to HAF/A4, balance the necessary input information with the required level of output fidelity, and ensure the applicability of the results to HAF/A4 concerns.

Process Extensions

To support novel insight into potential mitigations, Step 8 of the analysis was augmented to incorporate the following considerations relative to proposed controls:

- **Change to probability (threat)**: Many controls result in a change to the probability of successful threat action (e.g., alter the ability for the threat actor to mount a successful attack). Identifying the projected change to the threat enables the recalculation of risk scores and examination of what-if scenarios.
- **Change to consequence (impact)**: Other controls might change a risk's consequence in addition to (or instead of) threat probability. Examples include log audits, which do not

[144] Department of Defense Instruction 8510.01, *Risk Management Framework (RMF) for DoD Information Technology (IT)*, U.S. Department of Defense Office of the Chief Information Officer, July 19, 2022.

impede an intrusion but enhance the ability of an organization to respond and reduce the negative outcome.

- **Cost**: Ultimately, a primary consideration in the implementation of a new control is the cost of the control. Although comprehensive exploration of tool and manpower costs for a given control were out of the scope of this analysis, an estimated range (between one and three dollar signs [$], corresponding to the low-medium-high scale employed elsewhere) was provided to support the analysis.
- **Assumptions**: A critical aspect of informing cybersecurity decisionmaking is to state scope, efficacy, and cost (direct and indirect) assumptions explicitly. Assumptions employed in the calculation of risk inform the interpretation of the analysis.
- **Negatives**: As with assumptions, negative aspects of cybersecurity decisions should be considered when making control decisions. These aspects contribute to scope, cost, and efficacy estimation (capturing such aspects as opportunity cost or the impact a control might have on normal operation) and more fully develop the decision landscape.

Following the evaluation of controls individually, some effort was made to examine the impact of controls across risks in the context of the overall system.

- **Total risks addressed**: This captured the impact of each control against various risks. This is important to the cost-benefit trade-offs that are often made with respect to the scope of the intervention (how many individual risks are reduced by this control) versus the focus of the intervention (does this control target the highest risks).
- **Magnitude of risks addressed**: Complementary to the total risks addressed, the magnitude to which the risk was addressed (i.e., the effectiveness of the control) allows a decisionmaker to understand the extent to which the risk has been reduced (overall risk score) or the change in threat probability. The latter was calculated across each control-threat pairing (how well does this control mitigate this risk), across all risks (what is the overall threat reduction for every threat vector), and to determine average impact (which control provides the most benefit, on average).

Additionally, the execution of the process can (and, in many cases, should) occur following the addition of controls. This allows for an examination of how the threat surface changes in response to the addition of a control, as part of the cost-benefit analysis inherent in any cybersecurity decision.

Process Step Adaptation

In addition to the extensions and adaptations above, specific process steps were standardized to focus on the risk aspects most relevant to the HAF/A4 mission.

For Step 1, we consulted the previously cited HAF/A4 CIO Risk Frame to ensure that selected risk considerations were aligned to HAF/A4 CIO guidance.[145] Targeting *mission risk* as the appropriate level to capture IT and operational technology concerns and interactions for logistics, we

[145] USAF A4 (Program Integration Directorate), *A4 Chief Information Officer Cybersecurity Risk Frame*, September 4, 2020.

identified those risks that capture the time and scope concerns identified in prior RAND research.[146] This was augmented by an examination of the literature surrounding other such supply chain challenges, such as in the automotive supply chain.[147] Candidate measures were validated with internal DAF supply chain experts and selected for their relationship to mission effectiveness. This resulted in the identification of three measures that captured mission risk, shown in Table 4.1. A more directed evaluation of a specific system or scenario could easily replace these risks with other such project, mission, or organizational risks as identified in the Risk Frame to evaluate specific circumstances.

Table 4.1. Risk Measurement Criteria Employed for OCTAVE

Impact Category	Impact Area	Low	Moderate	High
Time criticality	Time to survive	Weeks to months	Days to weeks	Hours to days
	Time to recover	Hours to days	Days to weeks	Weeks to months
Scope of impact	Units affected	Platform	Wing	Fleet

SOURCES: Derived from Don Snyder, George E. Hart, Kristin F. Lynch, and John G. Drew, *Ensuring U.S. Air Force Operations During Cyber Attacks Against Combat Support Systems: Guidance for Where to Focus Mitigation Efforts*, RAND Corporation, RR-620-AF, 2015; and David Simchi-Levi, William Schmidt, Yehua Wei, Peter Yun Zhang, Keith Combs, Yao Ge, Oleg Gusikhin, Michael Sanders, and Don Zhang, "Identifying Risks and Mitigating Disruptions in the Automotive Supply Chain," *Interfaces*, Vol. 45, No. 5, October 2015; and RAND analysis.

As part of Step 2, internal SME workshops were employed to identify information assets for which a loss of integrity would hamper mission execution. Alternately, the process described in Snyder et al. (2015) provides a nonexpert, structured means by which potential information assets can be evaluated.

For Steps 4 and 5, we employed an analysis technique known as *misuse* (or *abuse*) *case development*.[148] This technique is most commonly associated with security requirement development and provides an attacker viewpoint on the system engineering concept of use cases. The open-ended, system-focused nature of the approach supported the goals of this analysis by enabling the same analysis approach for each system considered, using the information assets themselves as the core elements of the use cases.

Misuse cases were generated in two parts: a high-level diagram using standard unified modeling language and an accompanying structured account providing technical details that include attack course of events, alternate paths, pre- and post-conditions for the attack, and extension and exception points. Additional information (such as probabilities of attack or success and attacker cost) might

[146] Snyder et al., 2015.

[147] Simchi-Levi et al., 2015.

[148] Guttorm Sindre and Andreas L. Opdahl, "Eliciting Security Requirements with Misuse Cases," *Requirements Engineering*, Vol. 10, No. 1, January 2005.

also be captured to assist in later analysis.[149] The fleshed-out misuse cases provided the detailed information required to successfully complete the remaining OCTAVE Allegro steps. Although misuse cases were employed in this analysis, the information security literature includes several qualitative and quantitative approaches that may be employed in addition to, or instead of, this approach. For instance, the OCTAVE method proposes the use of generic attack trees.[150] Most importantly, techniques should be chosen to align with available information and the desired level of analytic detail.

To explore the question of mischief versus mayhem posed by HAF/A4 we enumerated and scored threats along two factors: capacity (the combination of skill and resources) and focus (the level of targeting involved in the attack). This simple threat model creates a set of four threat classes that employ variables captured in the misuse cases developed in Step 5. An additional threat class of *Insider* was added to represent concerns raised that that specific threat class. An evaluation of each was performed using the same low-medium-high scale employed for other risk calculations and aggregated into an overall threat rating. This construct easily allows reevaluation of risk scores along different individual threat actors or the ability to change threat actor values based on other information (e.g., intel sources, prior experience). The potential to conduct analysis across different scopes and types of threats and vulnerabilities is a key factor in OCTAVE's ability to support analysis across system types and levels of maturity.

In most scoring steps, we took a risk-seeking stance, intended to elevate the largest risks. This was accomplished by tracking the distribution of low-medium-high values assigned to each category and limiting the number of high values to be one-half of the number of low values.[151] This worked to frame risks and focus analysis on attacks and assets that pose the greatest risk (in this case, those with the largest impact on information integrity).

Finally, in Step 8 we sought to maintain consistency with existing processes by linking mitigations to recognized standards, such as National Institute of Standards and Technology (NIST) Special Publication 800-53 (revision 5),[152] the latest version of "Security and Privacy Controls for Information Systems and Organizations" employed within the RMF process. Candidate mitigations were examined against the risks along multiple security engineering dimensions. The goal of this linkage was to ensure that recommendations can be directly employed by a decisionmaker or incorporated into RMF analysis to support authority to operate (ATO) decisions.

---

[149] Chad Heitzenrater and Andrew Simpson, "Misuse, Abuse and Reuse: Economic Utility Functions for Characterising Security Requirements," in *Proceedings of the 2nd International Workshop on Agile Secure Development*, August 2016.

[150] In addition to the simple attack trees identified in OCTAVE Allegro, the development and analysis of attack trees has long been an active area of research within the cybersecurity literature. A good overview of approaches to attack tree modeling can be found in Barbara Kordy, Ludovic Pièetre-Cambacédès, and Patrick Schweitzer, "DAG-Based Attack and Defense Modeling: Don't Miss the Forest for the Attack Trees," *Computer Science Review*, Vols. 13–14, November 2014; and Adam Shostack, *Threat Modeling: Designing for Security*, 1st ed., Wiley Publishing, 2014. An approach to combine attack trees with misuse cases is discussed in Inger Anne Tøndel, Jostein Jensen, and Lillian Røstad, "Combining Misuse Cases with Attack Trees and Security Activity Models," *Proceedings of 2010 International Conference on Availability, Reliability and Security, ARES*, 2010.

[151] This is a practice that has been employed in other relative-risk methodologies to control the outcome risk profile. Participants may also take such stances as *risk-neutral* (using the same number of each risk rating) or *risk averse* in which risk values are weighted toward higher ratings, resulting in a greater number of items identified as high risk.

[152] NIST, *Security and Privacy Controls for Information Systems and Organizations*, SP 800-53, revision 5, September 2020.

Results

To demonstrate the utility of the process, we applied the process described above against three specific use cases of interest to the HAF/A4 mission (hereafter referred to as *threads*):

- *Thread 1: Demand forecasting*. This thread captures a general process flow related to the development of a demand forecast from disparate data sources, resulting in a processed output that feeds the ordering process.
- *Thread 2: Bot development and use*. In this thread, an idealized workflow for the development, testing, storage, and use of business applications using a low- or no-code platform is defined and analyzed. It does not focus on any specific bot use case but instead examines risk to bot development and deployment overall.
- *Thread 3: Bot use for data integration to enable failure analysis*. This thread builds on the concepts presented within Threads 1 and 2 to examine a specific bot use case, as described in Chapter 3. It envisions the operation of a bot with the task of compiling information from operational databases and producing an output record for failure analysis.

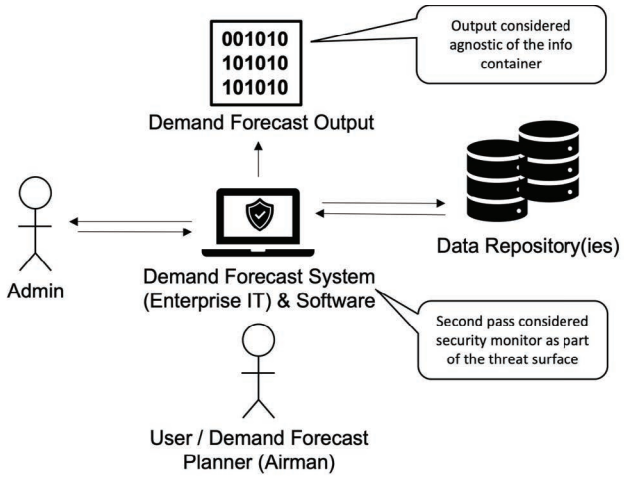
These threads were chosen for their links to the other aspects of this project. Thread 1 represents a use case taken from the discussion in Chapter 2, serving as an example of applying OCTAVE Allegro to current and evolving systems (i.e., ESCAPE). Threads 2 and 3 provide insights into the discussion in Chapter 3, demonstrating the use of OCTAVE Allegro to inform implementation decisionmaking in a technology adoption use case. Each thread is described graphically and in detail below, along with the results from our example analysis.

Our analysis of these threads demonstrates the application of a security framework for HAF/A4 and meets specific project goals: first, to provide a broad viewpoint of potential threats that is different from and complementary to the viewpoint supplied by the current, RMF-driven analysis already occurring for the as-is and to-be architectures. Second, our implementation of the OCTAVE was structured to minimize impact on existing processes. This was done in part to avoid affecting the HAF/A4P cybersecurity team, who could not support multiple workshops or enable access to all necessary materials. Our implementation demonstrates how HAF/A4 might conduct such an analysis without disrupting current security operations. Finally, the application of Thread 1 and Thread 2 outputs provides an example of how ongoing analysis can be employed to make recommendations relative to a future use case.

Thread 1: Demand Forecasting

Our first thread draws on current operations by examining the use of a system (demand forecast system) by a user/demand forecast planner (airman) employing software to develop a demand forecast output. This is accomplished by accessing one or more data repositories. Additionally, it is assumed that the system is administered by one or more administrative users. This is depicted in Figure 4.2.

Figure 4.2. Depiction of the Misuse Cases Identified as Part of Thread 1, Step 4



Drawing on domain expertise, we identified three primary information asset classes as part of Step 2 (e.g., flying hours, number of parts failures),[153] a forecast model algorithm (implementation), and the demand forecast (algorithm output).

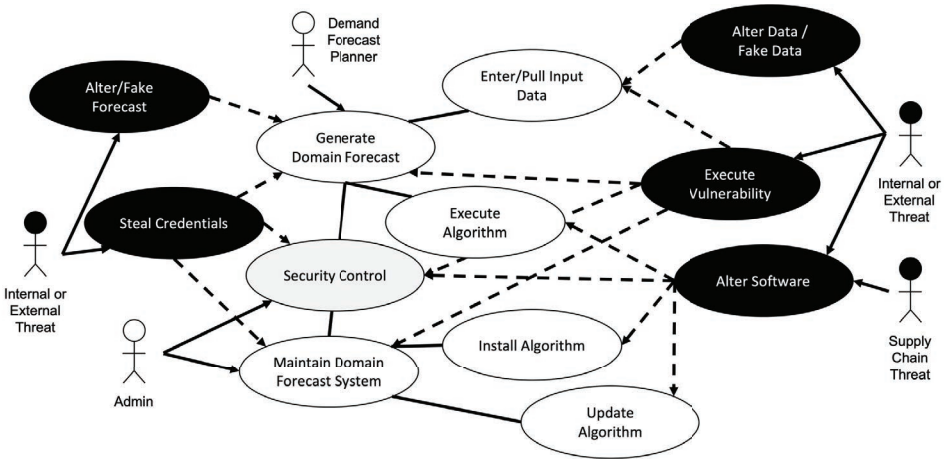
Each of these was evaluated relative to integrity concerns, resulting in the identification of five potential misuse cases. Figure 4.3 is a depiction of the potential misuse cases. The information assets and use case actions (white circles) employed in carrying out the scenario depicted in Figure 4.2 are affected by misuser or threat actor actions (black circles).

- *Altering/faking data* occurs when the misuser employs access to create, change, or otherwise manipulate data employed in demand forecast generation.
- *Altering/faking forecast* occurs when a false demand forecast is generated (using legitimate or illegitimate access) or an existing demand forecast is altered in a way that misrepresents the actual demand.
- *Altering software* occurs when the implementation of the demand forecasting algorithm is altered to produce incorrect results.
- *Executing vulnerability* occurs when a supporting misuse case permits misuser access to an information asset container.
- *Stealing credentials* is another supporting misuse case in which legitimate credentials are misappropriated or misused as part of another objective.

The *supporting* distinction refers to the objective of the misuser action, separating misuse cases that directly affect an information asset (e.g., alteration of the demand forecasting software) from misuse cases that could support multiple misuse cases (e.g., a vulnerability that permits access to the demand forecast software for alteration or other types of misuse).

[153] These were originally enumerated as individual information assets. As we conducted the analysis, we determined that, absent specific data relative to their generation and storage, differentiation did not contribute further insight. As a result, these were combined into *input data*.

Figure 4.3. Depiction of the Misuse Cases Identified as Part of Thread 1, Step 4



NOTE: The Security Control asset is gray to indicate its addition as part of a second pass of the analysis.

Applying these misuse cases to the three identified information assets resulted in 78 primary risks in Step 5. These were subsequently scored (Step 6) and then ranked according to the established risk criteria through a series of small and large group team discussions (Step 7), leading to a set of risk pools numbered 1 (highest) to 4 (lowest) (Step 8). After this initial scoring, an additional information asset was added. This generic *security control* represents the monitoring tools commonly employed on the Air Force Network (AFNet) and is depicted in the gray circle in Figure 4.3. Steps 2–7 were re-executed, resulting in a larger set of potential risks (83).

Risk and Mitigation Analysis Results

Determining the controls to implement for any mission set must consider factors that are not necessarily captured in this analysis, such as overall system security posture and system development concerns. This analysis focused on a core set of 20 common integrity controls applied to the risks identified in the upper risk pools (Pool 1 and Pool 2), allowing the research team to appropriately scope the analysis and identify key aspects important to decisionmaking:

- Mitigations with the broadest impact (i.e., contributed to the largest number of Pool 1 and Pool 2 risks) included those related to monitoring and event logging—both already implemented in some way as part of the system's current defensive posture. As reactive measures, these also had some of the lowest scores relative to risk reduction, having some impact on consequence but little contribution to the threat reduction component of risk.[154]
- Mitigations with the largest impact include cryptographic measures to protect the demand forecast and to validate the demand forecast software. Another high-impact (although likely unimplementable) control includes removing remote access to the domain forecast software. Although such a measure could have high costs in productivity or deployment options

[154] It should be noted that this is a matter of debate within the security community. While effective detection may conceivably halt an integrity attack, IBM cites the average time to identify a breach as 207 days—more than sufficient time to execute an integrity attack with mission impact (IBM, *Cost of a Data Breach Report 2022*, 2022).

(depending on architectural and operational considerations), the removal of access to the broader internet would mitigate many threat vectors to some of the highest-risk assets.

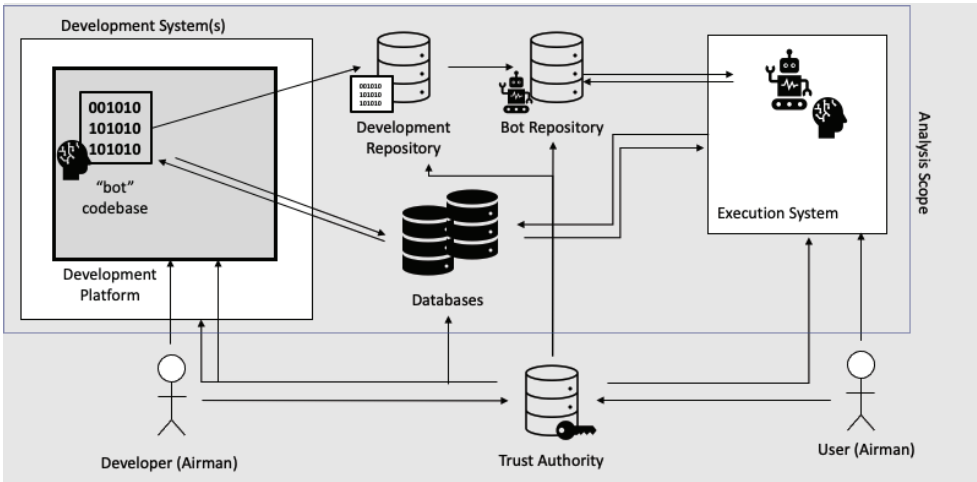
- The cryptographic mechanisms employed to manage and validate demand forecast software, output, and users are the most expensive. This is largely because of the required technical and procedural changes, which include management of these systems and the additional effort and training required. High-impact controls around software and code analysis will vary depending on the vendor's process maturity, expertise, and willingness to pass along these costs as part of the acquisition process.

The details of the risk and mitigation analysis behind these results, including a mapping of the considered mitigations to NIST SP 800-53 revision 5 controls, can be found in Appendix E.

Thread 2: Bot Development and Employment

Unlike Thread 1, Thread 2 describes an environment and mission set that has not yet been fully realized: a robust bot development and utilization environment, as described in Chapter 3. Figure 4.4 illustrates this scenario, which was employed to consider risks and mitigations prior to future bot development.

Figure 4.4. Notional Bot Development and Employment Scenario Used for Analysis

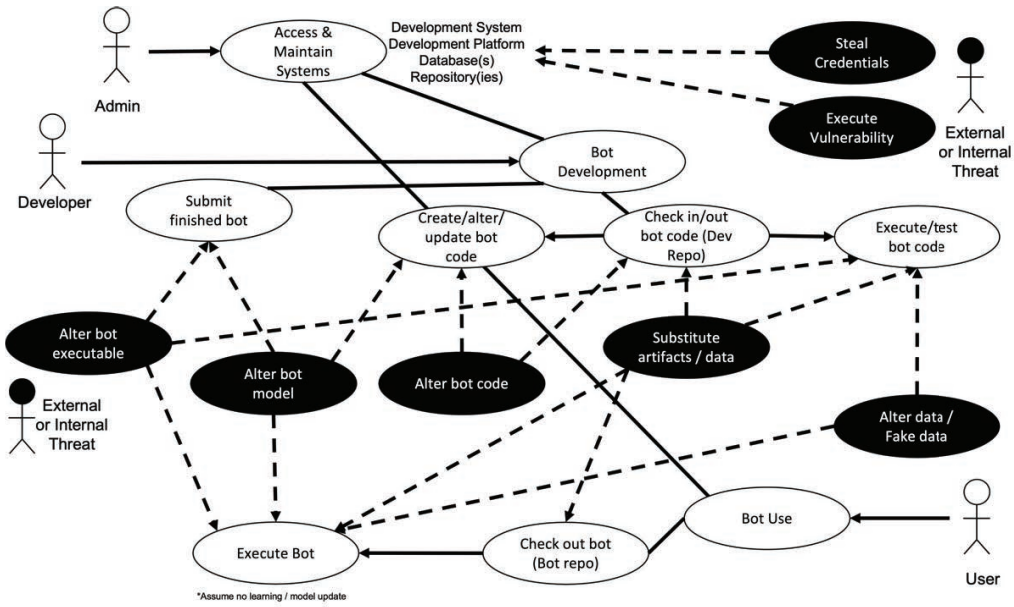


NOTE: This scenario assumes a static bot model and does not differentiate between operational and test data. The trust authority (e.g., a public key infrastructure [PKI]/CAC server) is not considered in the analysis.

The depicted scenario envisions a development system that hosts a development platform, on which the bot is developed and tested by a developer (airman)—potentially accessing databases to do so. The resulting code (and, potentially, the model) is stored in a development repository, which may be separate from (but connected to) the bot repository that stores compiled, ready-to-use bots. It is from the bot repository that the user (airman) accesses bots for employment, via an execution system. This scenario assumes static bot models (as opposed to learning bots) and explicitly recognizes that a trust authority (e.g., a CAC-enabled PKI) is present but might not be supported by all systems.

From this scenario, four information assets were identified: bot source code, bot (compiled), bot model, and data (test and operational). From this scenario, seven potential misuse cases were articulated, as depicted in Figure 4.5.

Figure 4.5. Misuse Diagram for Thread 2



NOTE: Dev = development; Repo = repository. White circles and solid arrows refer to user actions with or on information assets, while black circles and dashed arrows represent potential attacker actions against those assets.

As in Thread 1, the misuse cases identified represented a mixture of cases directly affecting information assets and those that support other misuse cases. In addition to the two supporting misuse cases already discussed (steal credentials and execute vulnerability), the following five misuse cases were defined:

- *Alter bot executable* occurs when compiled bot code is changed in a way that affects its execution or output.
- *Alter bot model* refers to the data or decision model employed by the bot executable. For most current bots, this model is represented in code as part of the system executable; however, a bot that incorporates ML may use ML that results in a model generated and stored separately as part of the development and training process.
- *Alter bot code* occurs when the code or commands employed in defining the bot are altered.
- *Substitute artifacts/data* refers to an instance in which a threat changes or substitutes artifacts (results, reports, diagrams) or data employed in the creation or execution of the bot.
- *Alter/fake data* occurs when the data employed in developing, training, or testing are altered or forged to undermine those processes.

Risk and Mitigation Analysis Results

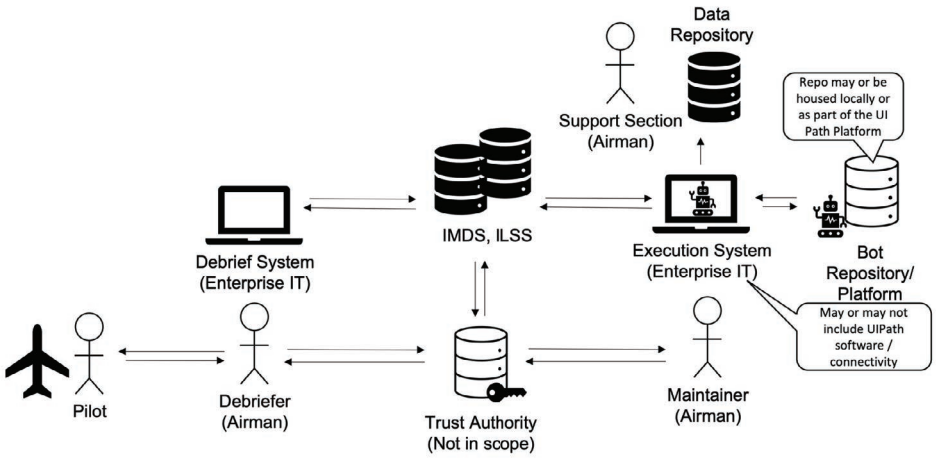
Risk and misuse analysis for this thread identified 70 risks, which were pooled and analyzed against the described scenario. Appendix E provides detailed findings that relate specific mitigations to NIST SP 800-53 revision 5 controls and "Open Web Application Security Project (OWASP) Top 10 Low-Code/No-Code (LCNC) Security Risks."[155] From this detailed analysis a few broad observations can be made for HAF/A4 consideration in the implementation of bot development:

- Remediation techniques (such as audits, standards, and policy) remain the mitigations with the broadest impact—although generally not the largest reduction in risk. In addition to audits, software security process and design practices (such as disabling user and service accounts) represent some of the lowest cost controls considered.
- Although the promise of low- and no-code environments is their utility to nonprogrammers, it will still be essential to maintain a level of quality in bot development operation to limit the introduction of vulnerabilities or errors leading to compromises of integrity.
- Some of the largest impacts to integrity risk, both per risk and across risks, involve efforts to secure the development environment itself.

Thread 3: Bot Employment for Data Integration to Enable Failure Analysis

For Thread 3, we sought to bring together the insights generated in Threads 1 and 2 to apply these concepts to the scenario outlined in Chapter 3 and depicted in Figure 4.6.

Figure 4.6. Depiction of the Thread 3 Bot Use Case



This scenario depicts a pilot providing a debrief to a debriefer (airman) who enters the information into a debrief system that stores details into the IMDS and ILS-S databases. A maintainer (airman) employs a bot (retrieved from a bot repository or platform) via an execution

[155] OWASP is an industry-recognized consortium that frequently issues lists of common flaws in an effort to raise awareness of security concerns (OWASP, "OWASP Top 10 Low-Code/No-Code Security Risks," webpage, undated).

system, which accesses the data stored in IMDS and ILS-S to produce an output stored in a data repository where a support section (airman) can access the output.

For this thread, we further scoped the bot aspect of the scenario by assuming the use of the UiPath platform and focused only on bots developed and applied to the failure analysis problem (analogous to the demand forecasting problem explored in Thread 1). The goal of this exercise was to demonstrate application of security analysis to a technology solution currently under consideration by HAF/A4, supporting investment decision processes.

Risk and Mitigation Analysis Results

Building on the existing analysis conducted via the previous threats, we identified 50 potential risks and several potential mitigation actions for HAF/A4 to consider in the implementation of a maintenance and supply data integration bot or for bots in general.

- The largest risks in our analysis remain those related to undermining the software or platform. This is especially relevant in RPA environments, which use significant amounts of vendor-controlled middleware and offer various configurations. Steps to take include the following:
  - *Ensure that processes for patching are in place.* It is important not only to consider the platform but also to include runtime environments (endpoints, browsers, or runtime libraries from the vendor) and dependent frameworks.[156]
  - *Maintain a security-focused configuration.* Ensuring the use of encryption for inter-process communication and data at rest (inputs, outputs, logs, etc.), disabling unnecessary features, and enabling authentication logs are all actions that an administrator should enact on the platform by default.
  - *Understand the security state of employed platforms.* Like all software, RPA environments have vulnerabilities, but the system security development life cycle (SSDLC) activities around review, scanning, and test processes reduce their likelihood and severity.[157] Having insight into these processes will assist HAF/A4 cyber professionals in evaluating the true risk of use and ensuring that proper mitigations are in place.
  - *Red-team development and testing environments.* Because generating insight into vendor SSDLCs is notoriously difficult, the next best option is a robust and ongoing program of red-teaming to evaluate the effectiveness of the activities above. Although only as effective as the team and rules of engagement allow, this is a key means for assessing both coverage and effectiveness of controls.

[156] In addition to the software itself, many platforms are built on top of frameworks, such as Microsoft's .NET. It is important to consider such dependencies as part of the threat surface and include them in maintenance efforts. This consideration can be enabled through the adoption of such constructs as a vendor-supplied software bill of materials (SBOM) (Cybersecurity and Infrastructure Security Agency, "Software Bill of Materials (SBOM)," webpage, undated).

[157] A simple search identified a handful of vulnerabilities in versions of the UiPath platform, some that are severe and allow remote code execution (National Vulnerability Database, National Institute of Standards and Technology, undated). Insights into vendor practices will help contextualize the likelihood of further severe vulnerabilities.

- Implement two-factor authentication (2FA) wherever possible. Although DoD uses 2FA by virtue of CACs for all AFNet machines, these protections might not always extend through vendor platforms accessed via browsers or off-premises servers. Using strong authentication and encrypted communications channels, potentially through an alternate 2FA implementation, would remove a host of Pool 2 risks related to data and code storage and movement.
- Well-developed policies and procedures contribute to risks across the board. Although it does not have the largest impact per risk, this set of mitigations could have both proactive contributions of reducing likelihood (such as the OWASP recommendations related to RPA secure design, coding standards, and artifact review) and reductions in risk consequence by defining essential reactive processes (such as when bots should be removed from service or how to respond to events). This is especially critical in the case of RPA development and use by nondevelopers, who might not understand or appreciate their role in reducing the introduction of responsibilities.

Finally, we note that the sheer number of ways in which RPA platforms can be deployed and used necessitates careful consideration on the part of HAF/A4. Deployment and use architectures dictate such elements as the use, location, and storage of logging; where, how, and to whom data are transmitted; and what additional software is used as part of the process runtime (including vendor-produced and third-party ecosystems, such as browsers). All of this contributes to the overall threat surface and drives many of the risk estimates discussed above. These concepts are a prime example of early-life cycle decisionmaking with profound impact to risk decisions and are echoed in the recommendations outlined in Chapter 5.

Limitations and Extensions

As discussed in Appendix D, the use of OCTAVE Allegro (or any risk-based cybersecurity analysis technique) brings a set of pros and cons, some of which relate to the decisions made throughout the analysis. For this analysis, the research team sought a broad look at a specific class of risks: those risks related to data asset integrity, their effect on spare parts supply, and the subsequent impact on risk to mission. This section outlines decisions made relative to this goal and highlights how this base analysis could be further expanded in scope and fidelity.

- *Limited workshops*. Ideally, OCTAVE Allegro steps are conducted as workshops with system owners, practitioners, and the larger set of stakeholders involved. With the burden on HAF/A4P a consideration, we sought to limit interaction with stakeholders. As developed, outputs were scoped to provide input to the RMF process, where security expertise already resides. Because RMF is primarily an *ex post* activity, our application of OCTAVE demonstrates a means to incorporate such security thinking throughout the system's life cycle, demonstrating analyses that are broader in scope (e.g., considering additional system interactions) or of greater depth (e.g., examining specific IT), but ideally requiring more interactive and directed stakeholder involvement.

- *Relative assessments*. OCTAVE Allegro is a relative risk assessment process. By anchoring risk measurements in mission risk, the implemented process supports mission assurance but is ultimately limited by the basis for comparison. A holistic analysis would consider a broader context, such as risks to confidentiality and availability and how risk changes throughout the spectrum of conflict. In addition, other nontechnical aspects, such as compliance and decision authority, were not incorporated as part of the RMF but are critical considerations when implementing a system security posture.
- *Generic threat model*. This analysis employed a threat actor model that combined scores across the capacity and focus continuum (plus insiders), as described. A more detailed analysis of specific systems might consider each of these threats individually and execute the analysis against only the pertinent threats (e.g., high-capacity and high-focus threats, such as advanced persistent threats) or drive scores based on supplied intelligence.
- *Risk stance*. In addition to employing a broad threat model, our analysis took a risk-averse stance in our scoring (rather than a risk-neutral or risk-seeking stance, which would have resulted in a greater number of high-risk elements). This was done to focus attention on the most-salient risks. A different stance might better capture HAF/A4's true operational environment or leadership viewpoint and might be used to examine how risks change under competition, crisis, and conflict.

Each of these decisions represents a potential departure point for further investigation that could sharpen, adapt, or expand on the example analysis as presented. Importantly, these elements represent the core contribution of a method with the flexibility to adapt and inform cybersecurity investment decisions, moving the practice beyond static checklists and toward life cycle security engineering.

Observations and Findings

The need to establish and maintain an effective cybersecurity posture, especially in dynamic environments with systems undergoing migration and refresh, often leads to complex and sometimes conflicting demands that require analysis and trade-offs to resolve. This section summarizes the key themes identified in our analysis, providing a more general view of these challenges as they pertain to the A4 community's mission.

- Each of the three threads considered here identified the software supply chain, software vulnerabilities, and credential-based attacks as potential vectors for integrity attacks.
  - It is possible that CAC will not support 2FA between all entities, depending on the deployment architecture. In this case, it is essential to understand authentication boundaries between system components.
  - Although the centralized PKI supplied by CAC enablement supports protections against most credential-based attacks, software supply chain and software vulnerability mitigation measures are less well supported.

- Many of the highest-scored risks were rooted in the software supply chain rather than in data alone. This is primarily because of existing DAF processes to deal with data issues (such as latency and consistency); these processes can compensate for minor errors before they compound. However, movement toward automation (e.g., widespread adoption of bots) could reduce such safeguards. This highlights the importance of adopting security analysis as part of a decisionmaking process.
- Thread 1 focused on the risks associated with a loss of integrity of data processing software, which has the potential to lead to many hard-to-detect tampering attacks. Methods for addressing this threat vary in cost, complexity, and effectiveness.
  - Placing requirements on software vendors might not incur immediate costs but could limit choice of available software or incur greater time and cost in contracting while still requiring government expertise to verify effectiveness.
  - Measures undertaken in house (e.g., attestation) could offer greater effectiveness and more control but are also likely to require additional expertise, process, and technical insight.
  - As demonstrated by the inclusion of the security control as part of the threat surface, all software—regardless of source—should be considered for both positive and negative contributions to the risk posture.
- Thread 2 demonstrated analysis of potential future operations. This thread identified several risks related to both development and operation of bots and underscored the need to incorporate training and best practice.
  - The bot development and operations platform should be considered a potential vector for vulnerability and actively mitigated through technical and nontechnical means (patching, red-teaming, etc.).
  - Threats to both the development and operations environment should be considered; the former provides risks that could be particularly hard to detect—especially for citizen airmen developers with little formal training in software.
  - In addition to malicious threats, nonmalicious action by citizen airmen developers is a risk that should not be overlooked. Best practice, training, and review should be employed to help mitigate this risk.
  - This is not to say that citizen airmen developers are a bad idea but that the benefit of their use should be weighed against the potential security risks for the development of a given bot.
- Thread 3 brought together elements of Thread 1 and Thread 2 to demonstrate how analysis can be employed to drive decisionmaking of technologies under consideration, incorporating cybersecurity into acquisition life cycle and mission-execution considerations. Conducting such analysis throughout the process can inform key deployment decisions (e.g., the configuration of UiPath), affecting risk, cost, and effectiveness.
- Across all three threads, robust security incorporates a mixture of controls from across the variety of mitigations. Current HAF/A4 RMF controls are primarily focused on detection

and remediation (as described in Thread 1) and are unlikely to be sufficient to mitigate the focused, tampering-based mischief attacks under consideration here.

It is unclear whether integrity-based attacks are the greatest overall risks to the HAF/A4 mission. The presented analysis provides insight into integrity attacks against key functions. More importantly, the process presented provides a mechanism to place those risks in the context of overall security concerns, such as confidentiality or availability attacks across the spectrum of conflict. In doing so, this research demonstrates how augmenting existing security analysis with a life cycle approach, such as OCTAVE Allegro, would allow HAF/A4 to understand these risks and methods for mitigation more fully, leading to better-informed cybersecurity decisionmaking.

Chapter 5

Recommendations

Recognizing the complexity of analyzing USAF supply chains as a whole, our analysis focused on three separate and specific topics that were of particular interest to HAF/Logistics Directorate (A4L): demand forecasting, RPA, and cyber integrity. Through the course of our research on methods to improve demand forecasting, how to apply RPA to improve supply chain effectiveness and efficiency, and how to mitigate vulnerabilities associated with cyber tampering, we identified intersections that loosely tied these areas together.

A review of legacy USAF demand forecasting methods, current efforts to modernize USAF supply chain planning via ESCAPE, and recent academic work on emerging methods in spare part demand forecasting highlighted that more-integrated datasets could enable the exploration of new methods for demand forecasting that incorporate more part demand, usage, and maintenance data. In this study, the research on demand forecasting and the research on bots was combined for an RPA bot that would facilitate analysis and be of use to demand forecasters by linking supply and maintenance data. Although we did not conduct analysis using these integrated data and cannot conclude that these data would improve demand forecasting, the availability of such data is a prerequisite for exploring the possibility, and it likely has benefits beyond demand forecasting. Finally, we applied the cyber integrity analysis methodology to our example bot that would perform this data linking to provide concrete examples of the types of risks and mitigations that such a methodology would uncover.

Although we connected these three research areas as described above, they are still distinct topics, each with a set of observations, findings, and recommendations. In the sections that follow, we summarize the findings already described in each chapter, followed by specific recommendations for USAF to consider.

Demand Forecasting

Major Observations and Findings

- Demand forecasting for spare parts is a topic that has received significant attention both inside and outside USAF for decades. There are a variety of methods to forecast spare parts demand, though there is not a one-size-fits-all approach that has been shown to be the best in all cases.
- Analysis of the sources of demand forecast error revealed primary drivers (i.e., parts for EOHs and low-demand parts) that personnel within the 448th SCMW who study the problem were mostly aware of but others across the broader A4 community might not be aware of.

- USAF is in the midst of a major change to the way it forecasts spare parts demand as it shifts to ESCAPE, which uses a best-in-class software product that, since the inception of its use, has shown improvements in DFA metrics.
- There is ongoing research both within DoD and across industry on the use of more sophisticated techniques, such as AI/ML, for demand forecasting, though additional research is needed prior to broad implementation for USAF. These methods will likely require better integration of data, including part demand, use factors, and maintenance policies.
- Demand forecasting is just one part of supply chain planning, and it is unclear whether DFA is resulting in increased aircraft downtime.

Recommendations

Drawing on a review of legacy demand forecasting in USAF, analysis of recent drivers of USAF forecast error, in-process changes to USAF demand forecasting, and relevant academic research, we offer the following recommendations:

- ***The 448th SCMW should maximize the potential benefits offered by the investment already made in ESCAPE.*** However, USAF should analyze the value of demand forecast improvements to supply chain performance prior to making additional investments. USAF is already investing a large amount of resources to transition its legacy spare parts planning system to a best-in-class commercial system via the ESCAPE program. That transition is still in the very early stages: FY 2022, the year this study was completed, is the first year ESCAPE is being used for demand forecasting. Although implementation of forecasting systems has not yet received much attention in the literature, there is some evidence that organizations do not take full advantage of the capabilities of their forecasting software.[158] Toward that end, USAF should do the following:
  - Continue to evaluate the impact of ESCAPE on DFA. As the analysis of DFA showed, there is no single driver of forecast error, and the literature review showed that there is no silver bullet, single solution for forecasting. Thus, analysis of forecast error drivers should continue as ESCAPE matures.
  - Expand the use of SPM built-in functionality to improve demand forecasts. SPM provides significant potential improvements across the pre-processing, processing, and post-processing phases of forecasting. Examples include experimenting with statistical versus causal techniques or using SEM forecasting techniques to improve depot forecasts.
  - Conduct analyses across different metrics to better understand the drivers of supply chain efficiency and effectiveness and whether demand forecasting is limiting efficiency. (DoD and USAF have already made significant progress in defining and tracking specific metrics related to supply chain efficiency.)
- ***If additional investment in forecast accuracy improvement is warranted, the 448th SCMW should target specific areas of improvement.*** Doing so will likely require improved data

[158] Boylan and Syntetos, 2009.

cleaning and integration, which could have additional benefits. The current sources of forecast error vary widely, and USAF should target specific areas where promising research is ongoing.

- One potential area for improvement is the expansion of causal models for predicting intermittent spare part demand. As discussed previously, academic literature has begun to show that incorporation of multiple installed base features into causal models has the potential to improve forecast accuracy, in particular with the inclusion of maintenance schedules. However, in most cases, researchers note the challenges associated with data collection and cleaning given the tendency for these datasets to be spread across many systems, tables, and organizations. This rings particularly true for USAF. The Air Force Science Advisory Board pointed out in 2011 that USAF supply chain management has inefficiencies due, in part, to the "large number of independent databases used by the USAF in tracking its aircraft, their configurations, their maintenance actions, and the parts used for maintenance."[159]
- Discussions with SMEs from RAND and USAF indicated that connecting supply and maintenance data could have significant benefits well beyond just demand forecasting for spare parts. For example, connecting these data could provide a more complete view for a particular part; enable measurement of time between replacements to validate distribution of failures, globally, and identify time-phased implications; and enable a more in-depth causal failure analysis that identifies the impacts of climate and mission profile. And better understanding of the relationship between inherent and induced failures by part could lead to better demand forecasting.

Application of Robotic Process Automation to Improve Supply Chain

Major Observations and Findings

- The USAF A4 community is in its early engagement with bots—developing them in generally stovepiped functional areas to automate manual tasks. However, this approach might not allow the A4 community to fully leverage the potential of bots.
- Questions remain about whether USAF personnel possess the technical expertise to fully leverage bot technology, and the data suggest that this concern is warranted.
- Developing and implementing bots needs to be weighed against the cybersecurity vulnerabilities they introduce.
- Unified direction and guidance regarding bots might help USAF in general and the A4 community in particular best maximize the potential of bots.

[159] USAF Scientific Advisory Board, *Sustaining Air Force Aging Aircraft into the 21st Century*, Department of the Air Force, August 1, 2011.

Recommendations

- ***Expand the application of bots within the A4 community.*** Additionally, broaden the solicitation for bot suggestions to organizations like HAF. Consider bots for such staff functions as collecting data for monthly review or reports. Expanding bots to conduct analysis within the A4 enterprise could result in opportunities that are more complex, sophisticated, and reflective of the art of the possible. We recommend the following bot concepts:
  - Give priority to the bot suggested in Chapter 3 that links IMDS and ILS-S data. Pockets of analysts within USAF are manually integrating IMDS, IPB, and ILS-S data to conduct failure analysis (which can inform demand forecasting), suggesting that the integrated dataset has value. The dataset created by the bot could inform causal analysis for failures of parts that goes beyond what is currently done. The utility of this type of analysis is being explored by a small team at AFLCMC that has merged these datasets manually to analyze failures by region to assess whether break rates vary because of differences in weather, location, flying versus ground time, and more. This bot could be further expanded to leverage AI/ML to process unstructured information, such as entries in the comments section for a JCN entry in IMDS. Such IA bots process large numbers of documents that come in varied forms, applying ML and sometimes NLP to extract important information. This incremental step above RPA could give USAF a chance to see what AI/ML can accomplish in a constrained environment.[160]
  - USAF could benefit from a bot to scrape data from maintenance systems and flying training systems to produce an optimized flying schedule at a flying wing. This bot would need to interact with optimization software, which might or might not leverage AI/ML, to apply sophisticated optimization algorithms to determine a weekly flying schedule that considers short- and long-term operational training requirements and short- and long-term maintenance requirements.
  - USAF could benefit from a bot to access data across logistics functional areas or scrape data to assemble information for status reviews, report submissions, and supply chain metric reviews.
- ***Work with the USAF lead for RPA to establish service-wide standards for development and management of bots and advocate for funding to support increased security measures.*** Decentralized efforts leverage different bot development processes, conform to varying cybersecurity standards, and cannot be shared across the enterprise. Creating centralized policies, processes, and data systems is necessary for proper management and tracking of bot deployment. We offer the following items for recommended inclusion in the service-wide standard guidance:

[160] We offer this suggested bot with a word of caution. The effectiveness of an IA bot is determined by the data it is trained on. If the data are not reliable or do not accurately reflect reality, the results of any IA will also be unreliable. Many of the bots being developed in USAF today will identify and fix misalignment between data systems. Until full faith can be placed in the accuracy of the data available, very little faith should be placed in AI/ML bots to inform strategic decisions.

  - The requirement to apply a taxonomy (such as the one described in Chapter 3 and Appendix B) to understand the primary characteristics of a bot before development begins. It is vital to fully consider all of the systems, users, programs, and documents that a bot might interact with—and how the bot interacts with them—before development and testing.
  - The requirement to perform a mission assurance–based risk analysis prior to proceeding with development of the bot, such as demonstrated in Chapter 4. Risk-based cyber analysis that spans the life cycle of development and deployment should be applied as part of the decisionmaking process when adopting any new technology, such as bots. Careful consideration of architecture, deployment, and configuration in the context of the mission can have a significant impact on the assumed cyber risk. The risk analysis could be informed by insight gained from applying a taxonomy recommended in Chapter 3 to understand the primary characteristics of a bot before development begins.
  - The requirement that bots are developed within a well-maintained and well-managed testing environment and not given access to actual data repositories until properly tested to improve security of data during bot development. The analysis in Chapter 4 highlights the data integrity risks posed by bots and their associated software supply chain. To overcome issues with data latency, it is recommended that bots do not output artifact documents that involve data that change rapidly. Instead, updates should be made in real time to data stored in the repository while adhering to least-privilege principles to manage risk. Additional bots might be designed to clean and standardize the format of incoming data so that it can be leveraged by all users, and the environment could be hosted (on-premises or in the cloud) to support increased interoperability and availability. However, implementation of these concepts comes with risks and requires analysis to understand and mitigate the potential vulnerabilities and consequences.
  - The institution of mechanisms to track user and bot actions separately by giving the bot separate access credentials from those of the user. Current RPA bots within USAF act as though they are the user, leveraging the user's access credentials and CAC. This makes it impossible for a distinction to be made between the actions of the user and the bot. In cases in which a bot behaves unexpectedly and damages data or systems, it will be difficult to identify whether it was the user or bot causing the damage.

Cyber Tampering

Major Observations and Findings

- Each of the three threads considered in our cyber analysis identified the software supply chain, software vulnerabilities, and credential-based attacks as potential vectors for integrity attacks.
- Many of the highest-scored risks were rooted in the software supply chain rather than in data alone. This is because of a few factors, one of which is the lack of DAF visibility into the code employed in logistics operations.

- Analysis of spare part demand forecast generation (Thread 1) focused on the risks associated with a loss of integrity of data processing software, which has the potential to lead to many hard-to-detect tampering attacks.
- Analysis of bot implementation in general (Thread 2) demonstrated the application of the OCTAVE Allegro assessment and identified several risks related to both development and operation of bots and underscored the need to incorporate training and best practices.
- This research combined elements of demand forecasting and causal data generated by a bot to demonstrate how the application of OCTAVE Allegro can be employed to understand cyber risks and influence decisions regarding technologies under consideration.
- Current HAF/A4 RMF controls are primarily focused on detection and are unlikely to be sufficient to mitigate the focused, tampering-based mischief attacks under consideration in Chapter 4.

Recommendations

- ***Consider the specific mitigation approaches for integrity attacks identified by this analysis by incorporating these considerations into existing cybersecurity investment processes (i.e., RMF). This should be accomplished by placing these options in the context of the system life cycle, cyber risk factors, and spectrum of conflict.*** This analysis focused on the question posed by HAF/A4 related to the dangers posed by tampering attacks. To address this question, this research provides both an approach to examining cybersecurity risk in context and application of that approach to examine tampering attacks. The threads presented in Chapter 4 applied the OCTAVE Allegro approach to illustrative cases to identify and prioritize cyber-tampering risks; however, a broader analysis is required to place these risks in context. In particular, the unique and dynamic nature of military operations and how they differ in competition and conflict requires focused, analysis-driven insight beyond traditional cybersecurity. Our mission risk criteria considered three elements: time to survive, time to recover, and units affected. The values for these metrics could be dramatically different in conflict as opposed to competition. Chosen mitigation actions should be viewed from a risk-reward perspective, accounting for scope, effectiveness, and cost.

  Example: The analysis in Chapter 4 highlighted integrity-based attacks to information assets across a variety of actors, examining risks that spanned systems and networks. Linking these assets to mission-centric metrics focused attention on risks and mitigations with the greatest potential to disrupt key operations under specific scenarios.

- ***Going forward, continue to evaluate cyber risks in context by implementing a process for considering how threats, vulnerabilities, and consequences to missions change as new systems, technologies, and information-handling methods are considered and implemented.*** Effective cyber posture must be developed and managed against organizational and technical concerns, considering the variety of operations from competition to conflict. Operating under the assumption that no system or collection of systems to support a mission can be made perfectly secure, USAF must continually evaluate cyber risk; assess it in context; and update

it relative to threats, vulnerabilities, and mission impact across systems, information assets, and mission objectives.

Example: Threads 2 and 3 of the analysis in Chapter 4 provide an example of how cross-functional cyber risk analysis processes can be employed and continuously analyzed to drive decisionmaking when adopting new technologies.

- ***Employ best practices for executing risk-based processes, including the following:***
  - Engage SMEs in workshop environments to better understand the value of information assets to the supply chain's support to operational mission execution. Given the complexity of USAF supply chain operations in competition and additional challenges introduced in conflict, SMEs responsible for executing supply chain operations will be best suited to identify the most critical information assets and assist in prioritizing risk-based security assessments that need to be accomplished.
  - Complement the current HAF/A4 RMF with a cross-functional approach. Life cycle-focused, risk-based security engineering processes (such as OCTAVE or similar processes) complement existing efforts and provide decisionmakers with a more comprehensive understanding of cyber risk to mission. Such processes enable the consideration of both proactive and reactive measures that span policy, technicality, and personnel actions and can identify mitigation paths earlier in the life cycle, where they can be more effective and less costly.

    Example: Threads 2 and 3 highlighted risk calculations that were dependent on the specific RPA platform, vendor security practices, and platform configuration, each affecting information asset exposure and threat surface. Input on risk throughout the life cycle of technology adoption can inform system design and implementation and improve the scope, effectiveness, and cost-benefit ratio of specific security decisions.

Appendix A

Annotated Bibliography of Select Demand Forecasting Research

Demand Forecasting Methods

Amirkolaii, K. Nemati, A. Baboli, M. K. Shahzad, and R. Tonadre, "Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by Using Artificial Intelligence (AI)," *International Federation of Automatic Control-PapersOnLine*, Vol. 50, No. 1, July 2017.

In this article, Amirkolaii et al. applied an AI method, NN, to forecast irregular demand and measured its accuracy to mean square error. The authors found that there was low forecast accuracy in preexisting methods for irregular demand patterns, and they were unable to forecast demand because of unpredictable use of aircraft spare parts. The authors recommend using NN with 1 part/1 feature or 1 part/multiple features for the most accurate demand forecast.

Atchley, Walter D., Dorothy M. Clark, Salvatore J. Culosi, Lori Dunch, Robert C. Kline, Thomas E. Lang, Randy L. Moller, Matthew R. Peterson, and Michael R. Pouy, *Lifecycle Forecasting Improvement: Causative Research and Item Introduction Phase*, Logistics Management Institute, Report DL920T1, November 2010.

Atchley et al. evaluates DLA inventory excess and shortfalls caused by demand forecasting and by policies for introducing new items. The authors found that forecast accuracy was not the only driver of inventory excess and shortfalls. For example, when new parts are introduced, these items most often have intermittent demand. As a result, services over-forecast new parts when they are first introduced. Rather than focus on DFA, Atchley et al. recommends improving the ability to forecast newly introduced items. The authors acknowledge that some shortfalls and excesses will always exist, but changes in operations might reduce their size.

Babai, M. Z., A. Tsadiras, and C. Papadopoulos, "On the Empirical Performance of Some New Neural Network Methods for Forecasting Intermittent Demand," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 31, No. 3, July 2020.

Babai, Tsadiras, and Papadopoulos evaluate the use of NN methods to forecast intermittent demand. As a benchmark, the authors deployed five alternative methods: single exponential smoothing, Croston's method, Syntetos-Boylan approximation, Willemain's method, and Zhou and Viswanathan's method. For the NN approach, the authors used a method presented in Gutierrez et al. (2008). Each of these methods was applied to a collection of airline company parts. The authors found that Willemain's method and Zhou and Viswanathan's method

performed the worst in inventory efficiency and the NN method performed second best to single exponential smoothing.

Bachman, Tovey C., Pamela J. Williams, Kristen M. Cheman, Jeffrey Curtis, and Robert Carroll, "PNG: Effective Inventory Control for Items with Highly Variable Demand," *Interfaces*, Vol. 46, No. 1, 2015.

In this article, Bachman et al. present PNG software, a tool developed by the authors to aid in decisionmaking for infrequent demand and frequent, highly variable demand. The software combines two methods, Peak Policy and Next Gen model (whose names combined are *PNG*). The software offers trade-offs in terms of wait time and on-hand inventory value. Within the software, the Peak Policy method provides a risk profile for a complete portfolio of parts. The Next Gen model uses a cost function to apply penalty factors, including backorders, inventory value, and annual buys. Similar to Peak Policy, Next Gen is only used for the population of parts. The software appeared effective at balancing costs while meeting the DLA director's stated objective for a 90 percent fill rate.

Baisariyev, M., A. Bakytzhanuly, Y. Serik, B. Mukhanova, M. Z. Babai, M. Tsakalerou, and C. T. Papadopoulos, "Demand Forecasting Methods for Spare Parts Logistics for Aviation: A Real-World Implementation of the Bootstrap Method," *Procedia Manufacturing*, Vol. 55, 2021.

Baisariyev et al. applies the Bootstrap method for forecasting spare part demand in commercial aviation. The Bootstrap method was selected from a taxonomy of forecasting methods. The primary methods used in aviation logistics included exponential smoothing, Croston's method, Syntetos and Boylan approximation, modified Croston's, and Bootstrap. Of these methods, the Bootstrap method had the best performance when compared using the *mean absolute deviation*, a scale-dependent accuracy metric. The Bootstrap method was found to provide accurate results with intermittent demand patterns but had lower performance with lumpy demand patterns.

Brown, B. B., and M. A. Geisler, *Analysis of the Demand Patterns for B-47 Airframe Parts at Air Base Level*, RAND Corporation, RM-1297, 1954.

In this 1954 research memo, Brown and Geisler assessed B-47 spare parts to observe possible Poisson distributions in demand. Demand for B-47 parts proved to be highly erratic, particularly for parts that were higher cost. The vast majority of items (458 of 470) analyzed had insufficient data to identify patterns in demand.

Brown, Bernice B., *Characteristics of Demand for Aircraft Spare Parts*, RAND Corporation, R-292, 1956.

Brown (1956) furthers analysis on B-47 airframe parts and identifies sporadic characteristics of demand. Most parts have low demand; one-third of parts have no demand, and three-fourths of parts have demand that is so low that using demand data will lead to bad predictions. The unpredictability of when demand occurs can be caused by future program changes, engineering changes, and changes to conditions of use. In addition to these factors, unpredictability can be attributed to challenges in analyzing demand, including a lack of data and a "lack of satisfactory systemization" of randomness in the demand for spare parts. To address these uncertainties,

Brown recommends an ample stock of low-unit-cost parts. For high-moving, high-cost parts, Brown recommends predicting demand with a probability distribution to determine stock. Alternatively, high-cost parts that are slow moving should be supplied as needed. Brown provides several recommendations to reduce the logistical burden of these high-cost, slow-moving parts, including reducing procurement lead times, resupply times, and repair cycle times.

DeFrank, Joshua D., "A Condition Based Maintenance Approach to Forecasting B-1 Aircraft Parts," Air Force Institute of Technology, March 3, 2017.

In this master's thesis, DeFrank applies condition-based maintenance to improve demand forecasting. When applied to the B-1 aircraft, the condition-based maintenance approach reduced demand forecasting error by 2.46 percent when compared with current approaches.

Guo, Feng, Jun Diao, Qiuhong Zhao, Dexin Wang, and Qiang Sun, "A Double-Level Combination Approach for Demand Forecasting of Repairable Airplane Spare Parts Based on Turnover Data," *Computers & Industrial Engineering*, Vol. 110, August 2017.

Guo et al. introduces a double-level combination forecast model to improve demand forecasting for repairable spare parts. The double-level model incorporates two types of combination models: low-level combination and top-level combination. The authors found that the double-level combination model performed accurately when compared with a single-level combination model and a single-level model.

Hyndman, Rob J., "Another Look at Forecast-Accuracy Metrics for Intermittent Demand," *Foresight*, No. 4, June 2006.

There are three types of measures commonly used to assess DFA for intermittent demand: scale-dependent metrics, percentage-error metrics, and relative error metrics. Each of these metrics faces unique challenges for accurately assessing intermittent demand forecasts. First, scale-dependent metrics cannot be compared with different demand series. Second, percentage-error metrics allow for comparisons, but their infinite measurements are not appropriate for intermittent demand. Last, relative error metrics are similarly not applicable for intermittent demand. In this paper, Hyndman introduces a fourth metric, scale-free error metrics. The mean absolute standard error allows for comparison between models and evaluation of the accuracy between demand series.

O'Neal, Thomas R., *Sortie-Based Aircraft Component Demand Rate to Predict Requirements*, Air Force Institute of Technology, March 2020.

In this master's thesis, O'Neal identifies a method for improving DFA by using sorties in place of flying hours. The author applied this method to the F-16 and B-52 using a modified Poisson process and found that demand forecast errors were reduced by 15 percent.

Posadas, Sergio, Carl M. Kruger, Catherine M. Beazley, Russell S. Salley, John A. Stephenson, Esther C. Thron, and Justin D. Ward, "Forecasting Parts Demand Using Service Data and Machine Learning," Logistics Management Institute, January 2020.

Posadas et al. assesses the application of ML methods for forecasting demand of aircraft parts managed by DLA. The effectiveness of ML begins with data, a common challenge for demand forecasting. To forecast maintenance demands, the authors used part consumption data. In this analysis, the authors focused on generator converter units for the F/A-18 E/F. Although the effectiveness of ML on demand forecasting is limited by data issues, some ML methods can decrease the impact of data scarcity.

Syntetos, Aris A., John E. Boylan, and J. D. Croston, "On the Categorization of Demand Patterns," *Journal of the Operational Research Society*, Vol. 56, No. 5, August 25, 2004.

In traditional demand forecasting, forecasters first categorize demand patterns. Next, analysts compare forecasting methods with error measures. In this paper, Syntetos, Boylan, and Croston recommended an analysis of the demand series as a first step in demand forecasting. This approach allows the demand patterns to be determined and defined by the forecasting methods' performance rather than categorized arbitrarily.

Van der Auweraer, Sarah, and Robert N. Boute, "Forecasting Spare Part Demand Using Service Maintenance Information," *International Journal of Production Economics*, Vol. 213, July 2019.

Most service maintenance parts have intermittent demand. Consequently, as availability increases, there is reduced investment in inventories. In this analysis, Van der Auweraer and Boute used the failures with the number of machines in the field to forecast demand and reduce inventory levels while meeting service needs. The authors found that while there is potential to improve forecasting, data availability presents a challenge.

Reviews of Forecasting Methods

Bacchetti, Andrea, and Nicola Saccani, "Spare Parts Classification and Demand Forecasting for Stock Control: Investigating the Gap Between Research and Practice," *Omega*, Vol. 40, No. 6, December 2012.

Bacchetti and Saccani used ten case studies to assess the gap between research and practice in spare parts management. Foundational to this analysis is a detailed review of spare parts classifications and forecasting methods. Through this analysis, the authors found that there was "little (if any) adoption of ad-hoc methods and techniques for spare parts management . . . the lack of integrated approaches . . . and a rather low level of awareness about how to perform managerial improvements" (p. 731).

Boylan, John E., and Aris A. Syntetos, "Spare Parts Management: A Review of Forecasting Research and Extensions," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 21, No. 3, November 12, 2009.

Boylan and Syntetos reviewed the development of demand forecasting for spare parts. They categorized three types of strategies for improving demand forecasts present in the literature: pre-processing, processing, and post-processing. Pre-processing methods include the demand pattern categorization between fast/slow and intermittent/lumpy. These methods evaluate the time between demand and the size of that demand. Processing methods to apply demand forecasting methods are the primary focus of demand forecasting research. Last, post-processing methods evaluate contributions made by adjustments made by a manager's discernment. The authors highlighted the need for research in pre- and post-processing improvements rather than the primary focus on processing methods.

De Gooijer, Jan G., and Rob J. Hyndman, "25 Years of Time Series Forecasting," *International Journal of Forecasting*, Vol. 22, No. 3, 2006.

In this review, de Gooijer and Hyndman presented research contributions to time series forecasting in the Institute of Forecasters. The authors categorized these contributions by the models used: exponential smoothing, autoregressive integrated moving average (ARIMA) models, seasonality, state space and structural models and the Kalman filler, nonlinear models, long memory, autoregressive conditional heteroskedasticity (ARCH)/generalized autoregressive conditional heteroskedasticity (GARCH) models, and count data forecasting. Finally, the authors presented research that combines these methods and the variety of accuracy measures used to assess demand forecast models.

Pinçe, Çerağ, Laura Turrini, and Joern Meissner, "Intermittent Demand Forecasting for Spare Parts: A Critical Review," *Omega*, Vol. 105, No. 1, July 2021.

In this article, Pinçe, Turrini, and Meissner review demand forecasting for spare parts with intermittent demand. Demand forecasting methods are not easily compared because of the variation in performance measures resulting in a lack of standardization to determine performance benchmarks. Furthermore, analyses for forecasting methods often omit significance tests when appraising methods. The lack of significant testing limits the objective comparisons that can be made between methods. Finally, the authors also found that the role of judgment in the demand forecasting process is understudied and provides a research opportunity to understand how managerial expertise can improve inventory levels.

Syntetos, Aris A., Mohamed Zied Babai, John Boylan, Stephan Kolassa, and Konstantinos Nikolopoulos, "Supply Chain Forecasting: Theory, Practice, Their Gap and the Future," *European Journal of Operational Research*, Vol. 252, No. 1, November 2015.

In this review, the authors state that a comprehensive review of supply chain management requires a full picture of supply chain echelons: manufacturers, wholesalers, retailers, and consumers. In addition to echelons, a supply chain analysis should include the products, locations, and periods associated with the supply chain. The authors reviewed literature for each

of these supply chain dimensions and evaluated the gaps between theory and practice. In this review, the authors found that practitioners and software manufacturers often lag behind theoretical advancements in each dimension.

Van der Auweraer, Sarah, Robert N. Boute, and Aris A. Syntetos, "Forecasting Spare Part Demand with Installed Base Information: A Review," *International Journal of Forecasting*, Vol. 35, No. 1, 2019.

Van der Auweraer, Boute, and Syntetos reviewed literature on demand forecasting methods that use usage location information. Although other approaches to demand forecasting use historical data, installed base methods incorporate current usage information that links machines and maintenance policies to determine spare part demand. Although these methods are understudied compared with methods that use historical data, the authors found that research on installed base methods has increased over the past decade; 40 percent of identified articles were published within the past five years.

Foundational Contributions to Inventory Management

Chenoweth, Mary E., Jeremy Arkes, and Nancy Y. Moore, *Best Practices in Developing Proactive Supply Strategies for Air Force Low-Demand Service Parts*, RAND Corporation, MG-858-AF, 2010.

USAF low-demand service parts present unique supply challenges. In their study, Chenoweth, Arkes, and Moore found that most USAF requisitioned parts (three in four of 60,000) had fewer than six annual requisitions. These low-demand parts often originated from USAF logistics centers, and more than half of low-demand parts originated from two USAF logistics center locations. The most-expensive parts are engines, turbines, and related components. The suppliers for low-demand parts are similarly concentrated. For example, eight of the top ten suppliers for all USAF parts are also present in the top ten suppliers for low-demand parts. These low-demand parts also have limited overlap with parts identified as MICAP or Awaiting Parts. These characteristics of low-demand parts make it difficult to improve supply chain strategies. Therefore, the authors recommended supply chain strategies in the design and production phases.

Hodges, James S., and Raymond A. Pyles, *Onward Through the Fog: Uncertainty and Management Adaptation in Systems Analysis and Design*, RAND Corporation, R-3760-AF/A/OSD, 1990.

Hodges and Pyles evaluates the current state of policy analysis and finds that previous analyses have critically omitted or mishandled uncertainty. There are two types of uncertainty that should be considered: state of the world uncertainty and statistical uncertainty. Although studies have included statistical uncertainty in their evaluations, statistical uncertainty is not often evaluated. One example of a study that incorporates state of the world uncertainty is Coupling Logistics Operations to Meet Uncertainty and the Threat.

Appendix B

Bot Taxonomy

Lebeuf and colleagues characterized a bot according to three primary dimensions: environment, intrinsic (how the bot functions), and interaction.[161] Within each of these dimensions, numerous attributes describe all the key information we need to know about a bot.

Environment Dimension

The environment dimension describes the surroundings in which the bot operates. This refers to the machine or server that the bot inhabits, the network or networks that the bot may access, and the data systems that the bot manipulates. The environment can be described by type, scope, closure, dynamism, predictability, permanence, and population, defined in Table B.1.

Table B.1. Bot Characteristics in the Environment Dimension

Characteristic	Potential Values	Description of Value
Type	Stand-alone	The bot is hosted independently but can access platforms in the same manner as users.
	Platform	The bot can be hosted independently or through the platform but accesses the platform without a user.
Scope	Bounded	The environment is limited with respect to its size.
	Unbounded	There are no limits on size.
Closure	Closed	Access to the environment is limited by restrictions on access.
	Open	An open environment allows free access without restrictions.
Dynamism	Static	All changes in the environment are the result of the bot's actions.
	Dynamic	A dynamic environment changes as a result of actions outside the bot's control.

[161] Carlene Lebeuf, *A Taxonomy of Software Bots: Towards a Deeper Understanding of Software Bot Characteristics*, master's thesis, University of Victoria, 2018; and Lebeuf et al., 2019.

Characteristic	Potential Values	Description of Value
Permanence	Episodic	The actions only temporarily affect the environment's state. After the current interaction, the environment returns to its previous state.
	Sequential	The actions permanently change the environment's state and could affect future actions.
Population	Singular	The bot is the only member of the population.
	Countable	The population can be reasonably counted.
	Uncountable	The population cannot be reasonably counted.
Diversity	Homogeneous	All members of the population are the same type.
	Heterogeneous	A heterogeneous population has a diverse set of inhabitants.

The *type attribute* describes the setting (often a machine or system) that the bot inhabits, participates in, or accesses. This can be *stand-alone*, where the bot is tied not to a specific platform but to a local machine. Otherwise, the type is *platform*, where the bot can be hosted on a server. Examples of platforms include the following:

- computer systems (networks, operating systems, and databases)
- social platforms (Facebook)
- ambient platforms where users use voice commands (Alexa and Siri).

Scope refers to the size of the bot's environment, which can be *bounded* when the environment is limited with respect to its size (such as the capacity of the data systems that the bot accesses). Scope can also be *unbounded* when a bot is scraping the internet (such as Google results) or may explore networks at will. *Closure* indicates who can access the bot's environment. Closure can be described as either *closed* when access to the environment is limited (such as accessing a restricted system or housed on a laptop with a login) or *open* when others freely access the environment (such as Google and news sources). *Dynamism* refers to the degree to which the bot's environment changes, which can be either *static*, where all the changes in the environment can be attributed to the bot's actions, or *dynamic*, where changes in the environment can also be attributed to other users or bots. For instance, if a bot produces a slide deck and that was the only thing that could change in the environment, the environment would be static. If other things can change as the result of others' actions (such as when a bot accesses a data system that is also used by others), the environment is dynamic.

Permanence refers to how long the effects of the bot's actions remain. This can be described as *episodic*, where the actions performed only in the environment temporarily affect the environment. When an episodic bot terminates, the environment returns to the same state as before the bot ran. Alternatively, *sequential* bots permanently change the environment and may affect future bot actions. This includes creating, moving, and changing files; sending emails; and manipulating data.

Finally, *population* describes the active entities (humans or bots) within the environment. Population can be further broken into cardinality and diversity. *Cardinality* refers to the number of other entities in the environment, which can be described as *singular* when the bot is the only member of the population, *countable* when the population can be reasonably counted, or *uncountable* when the population is so large that it cannot be reasonably counted. *Diversity* refers to the composition of the population, which is either *homogeneous* when all members of the population are the same type or *heterogeneous* when the environment includes diverse types of users and bots. A heterogeneous environment can refer to incorporation of both human users and bots, leveraging users from different roles, privileges, and purposes or including bots that perform different actions.

Intrinsic Dimension

The intrinsic dimension describes properties belonging to the bot itself, where the bot's developer has complete control over each attribute within. These attributes include knowledge, reasoning, adaptability, goal orientation, delegation, specialization, anthropomorphism, and life cycle. Several of these characteristics apply to bots that leverage AI. As illustrated in Table B.2, many of these attributes can be broken down further.

Table B.2. Bot Characteristics in the Intrinsic Dimension

Characteristic	Potential Values	Description of Value
Knowledge		
Memory	Long-term	The bot can store and access past events and actions.
	Short-term	The bot can temporarily store and access the current context, events, and actions.
	Future	The bot can store predictions of future events and actions.
Source	Encoded	The bot's knowledge is directly encoded by the programmer or creator.
	Supplied	The bot's knowledge is provided by something in its environment.
	Learned	The bot's knowledge is inferred from its environment.
Reasoning		
Mechanisms	Scripted	The bot responds to predefined stimuli with preprogrammed responses.
	Planning	The bot does not have a predefined script mapping inputs to outputs but instead makes decisions based on the situation and its current knowledge.
	Mixed	The bot uses a combination of planned and scripted reasoning mechanisms.
Agency	None	The bot has no agency if it requires an external party to approve actions before the bot can perform them.

Characteristic	Potential Values	Description of Value
	Complete	The bot does not require permission to carry out the tasks required to realize its goals. Complete agency is often described as autonomy.
Predictability	Stochastic	Results of the bot's reasoning mechanism appear as though they are random.
	Deterministic	Results of the bot's reasoning mechanism are the same when provided with the same inputs and conditions.
	Mixed	Some input types are stochastic while others are predictable.
Visibility	None	All of the bot's decisions or actions are hidden.
	Transparent	The bot's decisions or actions leave visible traces when the bot is not actively trying to make its processes visible.
	Visible	The bot actively works to make its decisions or actions visible. The bot creates additional artifacts for the sole purpose of providing visibility into its decisions or actions.
Reactivity	Synchronous	The bot responds at the same time or very shortly after the stimuli is perceived.
	Mixed	The bot uses a mixture of synchronous and asynchronous response times.
	Asynchronous	The bot responds to the stimuli after some time has passed.
Scheduling	Single tasked	The bot is single tasked if it can only handle one stimulus or task at a time.
	Multiple	The bot is capable of handling more than one task at once. Multi-task bots process tasks based on the order in which they arrive or according to some prioritization metric.
Adaptability		
Constraints	Constrained	The bot is able to adapt its dimensions, but it is restricted by scope, extent, or activity.
	Open	The bot is freely able to adapt its behavior.
Source	Internal	The bot's adaptation process is triggered from within the bot itself.
	External	The bot's adaptation process is triggered by something in the bot's environment.
Guidance	Undirected	The adaptation outcome is not directly shaped or influenced by a source. Undirected adaptation is guided from within (e.g., from its programming), and the bot itself controls the adaptation process.
	Directed	The bot's adaptation is directed if its adaptation outcome is shaped by the source's actions.
Goal orientation		

Characteristic	Potential Values	Description of Value
Complexity	Low	The bot performs simple tasks.
	High	The bot performs a complex workflow, drawing from multiple data sources or applying complicated analysis.
Criticality	Low	The bot's tasks are low risk, are less important, or entail no security concerns.
	High	The bot's tasks are high risk, are very important, or entail security risks.
Attainability	Achievable	The bot's goals have a well-defined end state that can be feasibly reached.
	Homoeostasis	The bot's goals are never reached.
Explicitness	Explicit	The bot's goals are clearly defined with no need for interpretation.
	Implicit	The bot's goals are ambiguous.
Source	Internal	The bot's goals come from the bot.
	External	The bot's goals come from external users, bots, or systems.
Delegation	None	The bot does not have the authority to act on behalf of others.
	Partial	The bot has authority to act on behalf of the user but does not pretend to be the user.
	Complete	The bot has the authority to both act on behalf of and pretend to be the user.
Specialization	Generalist bot	The bot supports a wide variety of tasks and directs users to the appropriate external resources.
	Specialist bot	The bot is designed to perform specific tasks in a limited domain.
Anthropomorphism		
Name	None	The bot has no name.
	Representative	The bot takes its name from the company or service it provides.
	Unique	The bot has been given its own identifiable name.
Embodiment	None	The bot has no visible form.
	Embodied	The bot has a visible form (e.g., logo, avatar).
Age	None	The bot has no age.
	Static	The bot has an age that does not increase.
	Dynamic	The bot has an age that increases over time.
Gender	None	The bot has no identifiable gender.
	Gender	The bot has an identifiable gender.
Ethnicity	None	The bot has no ethnicity.

Characteristic	Potential Values	Description of Value
	Ethnicity	The bot has an identifiable ethnicity.
Profession	None	The bot has no profession.
	Profession	The bot has an identifiable profession.
Personality	None	The bot has no personality.
	Personality	The bot was intentionally programmed with a personality.
Emotions	None	The bot has no visible emotions.
	Superficial	The bot only displays emotions when interacting with others but emotions have no effect on behavior.
	Logical	The bot displays emotions, and emotions have an impact on behavior.
Life cycle		
Lifespan	Terminating	The bot has a terminating lifespan if it eventually stops of its own accord.
	Transient	The bot is running sometimes and not other times.
	Continuous	The bot never self-terminates.
Creation	Human	The bot was created by a human.
	Bot	The bot was created by another bot.
	System	The bot was created by another system. For example, many bots in online games are spawned by a system.
Reproduction	None	The bot is unable to create other bots.
	Reproductive	The bot is able to create other bots.

The *knowledge* attribute describes what the bot understands or information the bot has immediately, which can be further described by its memory and source. A bot's memory can be characterized as *long-term* when the bot is able to store and access past events. This memory could include log files of past events that shape how the bot behaves in the future. A bot has *short-term* memory when it is able to temporarily store and access the current context only, often in short-term memory or a temporary file. A bot's memory is designated as *future* when it can store and access predictions of future events. The *source* attribute determines where the bot's knowledge originates. The source may be *encoded* if the bot's knowledge is programmed directly into the bot's code by the developer, *supplied* if the bot's knowledge is provided by something in its environment (like a file or data system), or *learned* if the bot's knowledge is inferred by surveying the environment.

The *reasoning* attribute refers to a bot's capacity to apply logic to achieve its goals. This broad but crucial category can be further broken into mechanisms, agency, predictability, visibility, reactivity, and scheduling. *Mechanisms* describe the way that a bot processes inputs or generates outputs. This processing can be *scripted* when the bot responds to a set of predefined stimuli with a corresponding set of preprogrammed responses, *planning* if it makes decisions based on the situation and its current knowledge, or *mixed* when the bot uses a combination of planned and scripted mechanisms.

Agency describes a bot's ability to perform its tasks without interference. For instance, a bot might be required to get clearance or permission from an external party. Therefore, agency is either *none* if the bot requires an external party to approve actions or *complete* if the bot does not require permission to carry out the tasks required to realize its goals. Complete agency is also called *autonomy*.

Predictability describes the degree to which the bot's actions are deterministic. Predictability can be characterized as *stochastic* where the results of the actions performed are random or follow a statistical distribution, *deterministic* where the results of the bot's actions can be fully predicted, or *uncertain* where the results of the bot's actions can be partially predicted. RPA bots are deterministic when there is no randomness in behavior (i.e., the same inputs will yield the same outputs every time).

Visibility signifies the degree to which a bot makes its decisions or actions visible to others. This can be described as *none* when all its actions are hidden, *transparent* if its actions leave visible traces when the bot is not actively trying to make its processes visible, or *visible* when the bot actively works to make its actions visible. Obviously, bots that lack visibility pose a threat to security because changes to the environment cannot be attributed to the bot. Many RPA bots are considered transparent, where a user can see the bot acting, but no log is created to document actions. A visible bot might create such additional artifacts as log files for the sole purpose of documenting changes to the environment. However, when data change often, artifacts can create latency problems if the information in the artifact becomes obsolete quickly. A better practice is to locate a bot and data within a cloud-based system to do linkages.

Reactivity refers to the time the bot takes to respond to stimuli, which can be *synchronous* if the bot responds at the same time (or very shortly after) the stimuli are perceived, *asynchronous* if the bot responds to the stimuli after some time has passed, or *mixed* if the bot uses a mixture of synchronous and asynchronous response times.

Scheduling describes the bot's strategy for dealing with multiple inputs or outputs that need to be reasoned about. A *single-tasked* bot can only handle one stimulus or task at a time, whereas a *multiple-tasked* bot is capable of handling more than one stimulus or task at once. Multiple-tasked bots can perform tasks based on the order in which they arrive or a prioritization metric.

Adaptability indicates a bot's ability to modify its own functionality at runtime, where an *adaptive* bot can change some behaviors at runtime and a *nonadaptive* cannot. Adaptation usually involves leveraging AI to survey and learn from the environment before deciding what to do. Deterministic RPA bots are nonadaptive. If a bot is adaptive, then it can be further described by constraints, source, and guidance.

The adaptation of *constrained* bots is restricted by scope or activity, providing safeguards to ensure the bot functions well within its core responsibilities. Alternatively, an *open* bot is freely able to adapt its behavior without restrictions. This is dangerous for several reasons, which will be discussed later in this appendix. *Source* refers to the origin of a bot's adaptation. This can be *internal* if the bot's adaptation process is triggered by programming within the bot itself or *external* if the bot's adaptation process is triggered by something in the bot's environment, such as the existence or configuration of files or data.

Guidance signifies whether the bot had some form of support during its adaptation process. Guidance can be *directed* when the bot's adaptation outcome is shaped by such a source as an external force (user or bot), algorithm, or document. Otherwise, the bot's guidance is *undirected*. Examples of directed adaption include *configuration*, in which an external force directly manipulates the bot's behaviors; *subscription*, in which an external force adds or removes behaviors; and *reinforcement*, which either rewards or punishes the outcome of the bot's adaptation to influence future adaptations.

Goal orientation can be loosely defined as a future state that the bot is working toward. This future state could include a finished analysis, a completed document, or a manipulated dataset. Goal orientation can be further broken down into complexity, criticality, attainability, explicitness, and source. Some of these categories are necessarily ambiguous because the taxonomy is meant to apply generically to all bots. *Complexity* refers to the complications surrounding a bot's goals, which can simply be labeled *low* or *high*. Low-complexity goals might include simple extraction of data, completion of a template, or submission of a requisition. High-complexity goals include sophisticated analyses involving multiple complex datasets.

Criticality represents the level of importance or urgency associated with the goal, which can also be labeled *low* or *high*. A bot with low criticality performs low-risk work that is not urgent and does not inform major decisions**,** whereas the goals of a highly critical bot are necessary to operations and inform pivotal decisions. *Attainability* describes the bot's ability to complete or achieve its goal. Attainability can either be *achievable*, where the goal can be feasibly met, or *homeostasis*, where the goal is never met. *Homeostasis* refers to bots that are running perpetually.

Explicitness indicates the degree to which the bot's goals are explicitly defined, which can be *implicit* when the goal is ambiguous or relies on judgment and *explicit* when the goal is clearly defined with no uncertainty or need for interpretation. *Source* describes what triggered a new instance of a goal behavior, which can be *internal* or *external*. *Internal* signifies that a bot's goals are programmed into it. *External* goals are provided to the bot from a user, another bot, or something else in the environment (e.g., a file).

Delegation refers to the bot's authority to act on behalf of others. This can be described as *none* when the bot does not have the authority to act on behalf of others, *partial* if the bot has the authority to act on behalf of the user but does not pretend to be the user, or *complete* if the bot has the authority to both act on behalf of and pretend to be the user. Many of the bots deployed within USAF are complete, where it is impossible to tell the difference between actions of the bot and the user.

Specialization is the degree to which the bot focuses on a specific area or task. A *generalist* bot supports a wide variety of tasks, while a *specialist* bot is designed to perform specific tasks in a limited domain. Most RPA bots are specialist bots, performing the same specific task repeatedly.

Anthropomorphism signifies the degree to which the bot is given such human-like traits as a name, age, gender, ethnicity, physical embodiment, profession, personality, and even emotions. Examples of anthropomorphized bots include Siri and Alexa, who were given voices of adult females. When a chat bot is given a physical avatar that interacts with users, the bot may be given a rough age, an ethnicity, a gender, etc. Bots that show emotions, such as bots found in some video games, typically leverage AI.

Life cycle refers to the various phases that the bot goes through in its life, which can be further broken into lifespan, creation, and reproduction. *Lifespan* is the length of time that the bot would have run if left completely alone. Lifespan can be *terminating* if the bot eventually stops of its own

accord, *transient* if it passes in and out of existence, or *continuous* if the bot never self-terminates. Bots that act on a schedule would be considered transient, and bots that constantly scrape data are continuous.

Creation describes the way in which the bot was brought to life, which can be *human* when a human developer creates the bot, *bot* when the bot was created by another bot, or *system* when the bot was created by another system. For example, many bots in online games are spawned by a system. Finally, the *reproduction* attribute describes the ability of a bot to spawn other bots, which can be labeled *none* or *reproductive*.

Interaction Dimension

The third dimension is interaction, which describes the rules governing how a bot engages with different entities in its environment. The interaction dimension can be broken into the following attributes: access, sense, act, communicate, initiative, robustness, and mobility. The potential values for these characteristics are described in Table B.3.

Table B.3. Bot Characteristics in the Interaction Dimension

Characteristic	Potential Values	Description of Value
Access	None	The bot may not access any of its environment.
	Partial	The bot may access a subset of its environment.
	Complete	The bot may access all of its environment.
Sense	Non-sensing	The bot does not perceive any external stimuli in its environment.
	Sensing	The bot perceives stimuli in its environment with sensors.
Act	Non-acting	The bot does not try to act on or make changes to its environment.
	Acting	The bot tries to act on or make changes to its environment.
Communicate		
Disposition	Antagonistic	The bot attempts to undermine others.
	Competitive	The bot acts in favor of its own self-interests.
	Indifferent	The bot is unaware of the needs of others.
	Cooperative	The bot coordinates actions with others.
	Benevolent	The bot helps others, even when contrary to its own goals.
Veracity	Untruthful	The bot deceives others.
	Mixed	The bot exhibits both deceiving and truthful behaviors.
	Truthful	The bot does not attempt to deceive others.

Characteristic	Potential Values	Description of Value
Cardinality	One-one	The bot interacts with one individual at a time.
	One-many	The bot is capable of interacting with multiple users simultaneously.
	Many-many	The bot is capable of interacting with many users while they are also interacting among themselves.
Directionality	One-way	The bot is capable of receiving inputs from or sending outputs to a user.
	Two-way	The bot is capable of both receiving inputs from and sending outputs to a user.
Directness	Indirect	The bot communicates with others through mediators, nonmessage interactions, or artifacts.
	Direct	The bot communicates with others through direct messages or requests.
Language capability	None	The bot is not able to use human language.
	Keywords	The bot communicates using short, preprogrammed phrases.
	Natural language	The bot is able to communicate with NLP (e.g., chatbots).
	Conversation	The bot can engage in meaningful dialogue.
Initiative	Reactive	The bot initiates actions in response to a specific stimulus.
	Proactive	The bot is proactive to control the situation rather than responding to stimuli.
Robustness		
Error prevention	Bot	The bot is responsible for preventing errors in the inputs it is receiving.
	User	The bot relies on users to prevent errors in inputs.
Error correction	Bot	The bot is responsible for correcting errors in the inputs it receives.
	User	The bot relies on the user to correct errors and provides the user with a way to correct information.

The *access* attribute simply refers to the degree of freedom given to the bot when accessing the systems, files, and data in its environment. Values for access are *none* if the bot is not allowed to access any of its environment, *partial* if the bot is allowed to access a subset of its environment, or *complete* if the bot is allowed to access all of its environment. Many bots are *restricted* when they can access a subset of a server or data system.

The *sense* attribute signifies the degree to which the bot can perceive environmental stimuli. A bot can either be *sensing* when it proactively seeks to perceive stimuli in its environment or *non-sensing* when it does not. Bots that wait for a user to hit *run* are non-sensing. Examples of sensing bots

include bots that clean data when the data enter a repository or bots that document purchases after they are submitted.

The *act* attribute represents the bot's ability to act on or make changes in its environment. An *acting* bot makes changes to its environment, creating, moving, or manipulating files or data, while a *non-acting* bot does not.

Generally associated with AI bots, the *communicate* attribute refers to the degree to which the bot can have meaningful interactions with others. A *communicative* bot interacts with others in its environment, and a *noncommunicative* bot does not. The communicate attribute can be further broken down into disposition, veracity, cardinality, directionality, directness, and language capability.

Disposition describes a bot's willingness to help, perform actions for, or share resources with others in its environment. A bot's disposition could be *antagonistic* if it purposefully inconveniences others; *competitive* if it acts toward its own self-interests; *indifferent* if it is unaware of the needs of others (inadvertently or by choice); *cooperative* if it coordinates its efforts with others, or *benevolent* if it always helps others in its environment, even if doing so is detrimental to its own goals or best interests.

Veracity refers to how truthful the bot is during communications. An *untruthful* bot intentionally deceives, a *truthful* bot does not deceive, and a bot that is *mixed* is both deceptive and truthful.

Cardinality in the interaction dimension represents the number of users that the bot can interact with simultaneously. Cardinality can be designated *one-one* if the bot can only interact with one individual at a time, *one-many* if the bot can interact with many users at the same time, and *many-many* if it can interact with many users while they also interact among themselves.

Directionality determines whether the bot can receive inputs or send outputs to a user. *One-way* directionality means a bot is only capable of receiving inputs from or sending outputs to a user, but not both. *Two-way* directionality means it is capable of both receiving inputs from and sending outputs to a user. *Directness* describes the way the bot communicates with others in its environment, where a bot is described as *indirect* when it communicates through mediators or nonmessage interactions and *direct* when it communicates through direct messages or requests.

Language capability signifies the bot's ability to communicate using human language, where values may be *none*, *keywords*, *natural language*, or *conversation*. *Initiative* is related to the bot's ability to sense stimuli in its environment, reason about the changes it detects, and act. A bot's initiative can be described as *reactive* when it initiates actions in response to a specific trigger in its environment or *proactive* when it takes action to control the situation.

Robustness signifies how well a bot handles errors, which can be further broken into error prevention, error discovery, and error correction. *Error prevention* includes strategies that a bot uses to reduce or prevent errors when receiving inputs from users. These strategies can be described as *bot* when the bot is responsible for preventing errors in the inputs it is receiving and *user* when the bot relies on users to implement preventative processes. *Error discovery* describes strategies that the bot uses to detect errors in the inputs it has received, where values are also *bot* when the bot is responsible for discovering errors in the inputs it receives and *user* when the bot relies on users to detect errors. *Error correction* refers to the strategies that the bot uses to recover from detected errors in the inputs the bot has received, where values can be *bot* when the bot is responsible for correcting errors in the inputs it receives and *user* when the bot relies on the user to correct errors.

Summary

By using the taxonomy during the course of bot design, USAF can better understand and document the primary characteristics of a bot. Although most developers have a process for describing a bot prior to its development, a full taxonomy can raise questions for the bot's use that might not have been previously considered. Without a full taxonomy, HAF/ A4L could have an incomplete picture of the impact a bot will have on systems and other users, which could cause conflicts with other systems or could introduce additional risks. The information gathered through the process of applying this taxonomy could inform cyber vulnerability analysis highlighted in Chapter 4.

Appendix C

Cyber Tampering Analysis Data

Accompanying this report are the worksheets used in the analysis of the three threads in Chapter 4. This appendix describes these worksheets and provides necessary background to understand the OCTAVE Allegro process as employed. These workbooks were developed and customized for the augmented version of the process, with the input and output from each step linked to feed the requisite input values for other steps.

For each workbook, the naming convention employed consisted of the OCTAVE Step ("SX"), a description of the content, and the related OCTAVE worksheet (numbered 1 to 10). Threads 1 and 2 (AllegroWorksheets-Thread1-FINAL.xlsx and AllegroWorksheets-Thread2-FINAL.xlsx) contain each of the worksheets below. Thread 3 (AllegroWorksheets-Thread3-FINAL.xlsx) drew from these threads, resulting in minor changes noted in the sections below.

1. **Sheet 1** [***S1 Risk Mgmt, Priority (1–7)***]**.** This sheet captures the stakeholder inputs to the identified mission risk categories and space for stakeholder rating. This sheet is the same for each thread.
2. **Sheet 2** [***S2 Crit Asset Summary (8)***]**.** The purpose of this sheet is to specify and detail the information assets that form the basis for the analysis. The following fields were used:
    a. Asset: The information asset (by name)
    b. Rationale for Selection: The reason behind identifying that asset
    c. Description: More-detailed information on the nature of the asset
    d. Owner: The owner of the asset; for our analysis, this was largely deemed not applicable to the analysis
    e. Security Requirements: Captures the stakeholder evaluation of the relative importance of each security requirement: confidentiality (C), integrity (I), availability (A). For these analyses, each information asset was rated highest for integrity ("I") because of the focus on tampering attacks. Additionally, the team sought a risk-neutral stance with an equal number of high and low ratings among the assets.
    f. Most Important Requirement/Rationale: A brief explanation supporting the security focus.
3. **Sheet 3** [***S3 Risk Env Maps (9a, b, c)***]**.** A listing of the information containers (technical, physical, and people), internally and externally, involved with the information life cycle of the information assets listed on Sheet 2. A notes column was added to supply additional details where useful to the execution of the analysis.
4. **Sheet 4** [***S4–6 Misuse Summary***]**.** This worksheet summarizes the misuse case diagrams and misuse descriptions to begin the process of creating candidate risks for analysis, covering aspects of Steps 4, 5, and 6 in the process. The sheet contains the following columns:

a. Case: A unique identifier for each information asset-misuse actor-misuse path combination; used for reference
b. Information Asset: The information asset affected by the misuse case
c. Description: The misuse case description
d. Dependency: Any misuse case (or, in the case of insiders, threat actor) that enables the primary misuse case
e. Path: Variations (extensions) on the primary or dependency misuse case that present variations
f. Justification/Assumptions: Supporting information related to the potential risk
g. Consequence: The results of the misuse case, articulated in the language related to the risks identified in Step 1 (Sheet 1).

Items that mapped to preexisting misuse cases were removed from consideration when the risk tables were developed (Step 5). These items have a gray background.

5. **Sheet 5 [*S58 Info Asset Summary (10)*].** Complementing Sheet 4, this portion of the analysis supports Steps 5, 6, and 8 by linking the potential risks with their threat and consequence, leading to the development of the risk pools and supporting analysis of mitigations. The following fields are employed:

    a. Information Asset: The information asset related to this risk
    b. Area of Concern: The misuse case
    c. Means: The combinations of dependencies and paths for a given misuse case
    d. Security Reqts: The security concern affected (C, I, A). For each misuse case and information asset, only integrity ("I") cases were considered in this analysis.
    e. Threat Rankings/Threat Totals: The next four columns represent the stakeholder assessment of risk against a set of actors:
        i. *Low capability—Low focus:* Actors that constitute attack "noise," who do not possess significant means or are not targeting the A4 community.
        ii. *Mixed capability and skill:* Actors that may possess either skill or focus, but not both. This could represent criminal or ideological elements.
        iii. *High capability—High Focus:* This is the traditional "advanced persistent threat" threat actor, with both means and the intent to attack the A4 community.
        iv. *Insider:* This considers the unique probabilities that an actor without authorization but with virtual or physical access might enjoy. Note that this can include such nonmalicious intention as accidental data deletion.

        These values are summarized into an overall threat level, which is segmented into low-medium-high relative risk. The totals for each assignment of probability were tracked in the total to follow risk profiles, as described in Chapter 4.
    f. Consequence Rank, Threat × Consequence, and Risk Pool: The values resulting from the analysis of Sheet 6 (Step 7) are linked here, allowing the calculation of the overall risk value (threat × consequence) and assignment into the risk pool.
    g. Concerns: Any assumptions or information useful to the interpretation of the values assigned was recorded here.

h. Notes: External references and commentary captured during the analysis. For Thread 3, analysis of risk developed in Threads 1 and 2 was employed to drive the analysis. Therefore, the version of this sheet for Thread 3 focuses on mapping the assets and misuse cases between threads to drive the selection of risks in the next step.

6. Sheet 6 [*S7 Risk Analysis*]. Step 7 of the process brings together the risk measurement criteria identified in Step 1 (Sheet 1) with the consequence information articulated during Step 6 (Sheets 4 and 5) to derive the consequence values. These values are then employed in Sheet 6 as part of the risk value calculation and pooling. Sheet 6 captures the following:

    a. Information Asset: In OCTAVE Allegro, risk consequence is driven by information asset.
    b. Consequence: A description of the consequence resulting from the undermining of the identified risk metrics. This description should relate to the risk measurements identified in Step 1 and the misuse case arguments made in Step 4.
    c. Risk Score: Each information asset is scored high-medium-low against the risk considerations identified in Step 1, weighted according to their ranking. This results in an overall risk score. Once again, the number of each relative ranking is tracked to provide the analyst with insight into the risk posture; our analysis sought a risk-neutral posture for this step.

    Thread 3 employed risks derived from the analysis conducted in Threads 1 and 2.

7. **Sheet 7 [*S8 Mitigation Strategy*].** This complex sheet captures key elements generated on other sheets, serving as the primary outcome of the analysis:

    a. Risk Pool: Establishes values for segregating risks into pools for consideration
    b. Mitigation Approach: Identifies the risk action to be taken for each risk. The same approach matrix was employed for each thread, with mitigation actions only identified for risks in Pools 1 and 2 (to limit the scope of the analysis).
    c. Pool 1, Pool 2/High, and Pool 2/Low risks: Risks in Pools 1 and 2 are listed, along with the number and percentage for each risk pool. Relevant asset, threat, and path are listed for each risk. In addition:
        v. Potential mitigations that might apply to each risk are listed by letter. A table on the left of the sheet ("Mitigations") lists the total risks of each mitigation.
        vi. The associated row number from Sheet 6 is listed, for reference.
    d. Mitigation Effects: To support analysis of each risk and mitigation pairing, a table to the right of the risk pools lists anticipated effects of each paired mitigation (listed by letter) on the residual risk ("R") and the change in probability of the risk ("L") for each mitigation mapped to that risk. These are summarized at the top by listing the average and total probability reduction for each potential mitigation.

8. **Sheet 8 [*S8 Mitigation Approaches*].** The last portion of the analysis supports the identification and description of the mitigations used in Sheet 7. This table lists identified potential mitigations, assigning a letter (A–Z) to each, along with the following:

a. Control ID: The NIST SP 800-53 rev. 5 control associated with the mitigation. This is not exhaustive; some controls overlap.
b. Mitigation: A description of the mitigation
c. Mitigation Type: An indication of the nature of the mitigation (e.g., technical, procedural)
d. Information Asset Affected: A list of the information asset that *may* be affected by the proposed mitigation
e. Cons Score Change: Any change in the consequence score brought on by the mitigation
f. Path(s) Affected: The misuse case paths affected by the proposed mitigation.
g. Residual Risk, Threat Score, and Projected Residual Risk: These columns mirror the scoring columns in Sheet 6 but are intended to capture changes based on the mitigation by incorporating the "Cons Score Change" and an updated threat score. Note that this is listed as "no higher than" because, for some risks, mitigations may not reduce these values beyond their originally derived values.
h. Cost: This column captures a very rough cost estimate equivalent to high-medium-low ($$$, $$, $). This is not a precise estimate but is intended to help decisionmakers differentiate between more and less expensive mitigation options.
i. Negatives: All mitigations come with trade offs (including some amount of increased cost). This column captures any negatives identified as part of the analysis to support decisionmaker considerations.
j. Assumptions: Most mitigations in this analysis were described at a high level to support the intention of evaluating a class of attacks, leaving room for variations in implementation. Details that would affect the scoring of mitigations were recorded here.

9. **Sheet 9** [***Data***]**.** This sheet contains the data used to generate the validated lists employed in other sheets.

Appendix D

OCTAVE Allegro Details

This appendix provides further information on the rationale and motivation behind the selection of OCTAVE Allegro and details some of the benefits and limitations discussed in Chapter 4.

Comparison with Other Approaches

As is the case in security engineering generally, the practice of threat modeling is an evolving space with active research. As cybersecurity approaches continue to mature, approaches in this space tend to have three different points of emphasis:[162]

- *Focus on assets*. Asset-focused methodologies focus on attacker targets and the means by which those targets are accessed. Methods in this category include enumeration of sensitive information (e.g., passwords, personal information) and the application of concepts like the *CIA triad* (confidentiality, integrity, availability) to identify concerns.
- *Focus on attacks*. Attacker-focused approaches are popular because of their tangible nature, outlining specific actions. Outlining specific actions can be done using personas (idealized attackers) or, as is increasingly common, using observed attack methods or known intelligence. Although these methods benefit from their realism, they can be retrospective and limiting when considering broad issues. Microsoft's STRIDE (spoofing, tampering, repudiation, information disclosure, denial of service, escalation of privilege)[163] pneumonic is a popular method to focus on types of attack.
- *Focus on systems*. Systems-focused threat modeling uses knowledge of the software or system to identify avenues of attack, which can be combined with attacker profiles to create detailed descriptions of how an attack might unfold. Attack trees,[164] which share a common lineage with fault trees and Byzantine fault analysis from the safety discipline, are common methods but require detailed information that might not be available—especially in the case of systems not yet designed or developed.

Clearly, limiting the scope of our analysis to one of these approaches would not have met the goal of providing a mission-centric, cross-functional view of the A4 community's cybersecurity needs. As a result, we sought to incorporate asset, threat, and system viewpoints into a broader picture, while

[162] Shostack, 2014.

[163] This method is described in more detail by Adam Shostack and the Microsoft Security Team on its blog (Adam Shostack, "STRIDE Chart," *Microsoft Security* blog, September 11, 2007).

[164] The concept of attack trees were first described by noted security author Bruce Schneier in a blog post, the computer security community adopted the concept and developed a body of literature that now promotes multiple published methods and analysis techniques (Bruce Schneier, "Attack Trees," *Schneier on Security* blog, December 1999).

making the best use of the available data and interaction we were able to have with HAF/A4 cyber personnel.

Why OCTAVE?

It has been long recognized that threat modeling of operations larger than a specific system or piece of software in ways that are predictable, are consistent, and offer a high ROI requires a dedicated process. OCTAVE Allegro has been recognized by the community as a mature means for modeling at this level.[165] Aspects of OCTAVE lend themselves well to the challenge of evaluating the myriad integrity concerns within logistics systems, including the following:

- *Abstraction.* As a risk-centric approach, OCTAVE can be run at various levels of abstraction.[166] As a methodology that was originally conceived as part of software and system development, both realized (implemented) and designed systems can be evaluated using the same process if the system can be sufficiently described and risk measurements applied.
- *Application of measures.* Although prescriptive to the goals at each step, OCTAVE does not dictate the measurements or approaches to meeting process goals. This allows for the incorporation of both quantitative and qualitative approaches (such as modeling and simulation, intelligence, or data analytics) as part of the analysis. The ability to apply varied methods contributes to the flexibility described above.
- *Documentation.* Although the approach might vary, the output at each step of OCTAVE is well defined and leads to standardized results that lend themselves to further analysis, reevaluation, or incorporation into other processes.
- *Relationship to HAF/A4 processes (e.g., RMF).* This flexibility within a standardized framework was used to incorporate materials provided by HAF/A4 (such as the HAF/A4 CIO Risk Frame)[167] into the process, while also providing output that could be employed to further RMF and ATO decisions.

One way to think about these aspects is to think of OCTAVE Allegro like a generalized planning process for security. Just as plans can be made at various levels of fidelity (or abstraction) and subsequently evaluated along different measure dimensions depending on the goal (such as efficiency or cost), OCTAVE and similar constructs support security decisionmaking throughout the system life cycle. Where RMF evaluates the state of security post-deployment, OCTAVE supports an understanding of how well security meets needs—similar to verification and validation in planning.

As with any analysis, there were both pros and cons to this approach. On the positive side, the information-centric approach (employing information assets as the primary focus of the analysis) provided a tangible link between the systems and infrastructure (where information exists) and the

[165] Shostack, 2014.

[166] The term *abstraction* is used here relative to its definition in computer science: "The act or process of leaving out of consideration one or more properties of a complex object so as to attend to others" (Jeff Kramer, "Is Abstraction the Key to Computing?" *Communications of the Association for Computing Machinery*, Vol. 50, No. 4, April 2007). In this case, abstraction allows the analysis to focus on different aspects and levels of risk.

[167] U.S. Air Force A4P, 2020.

mission (how the information is used). This, in turn, allowed for a mission-centric, infrastructure-agnostic view, tying recommended mitigations to specific mission risk through identified critical data elements. Being risk based, the output of this process (in the form of recommendations on actions or controls) can then be integrated into such standard processes as RMF to support investment, authorization, or compliance decisions. However, the abstract and flexible nature of the process requires careful interpretation and care at each step. Otherwise, the scope of considerations could grow exponentially or lead to analysis that inappropriately identifies large numbers of risks as high (or low), leading to a lack of differentiation and insight. Grounding the execution of the process through RMF and best practices helps to ensure that the result is usable and contributes to HAF/A4's cyber posture.

Appendix E

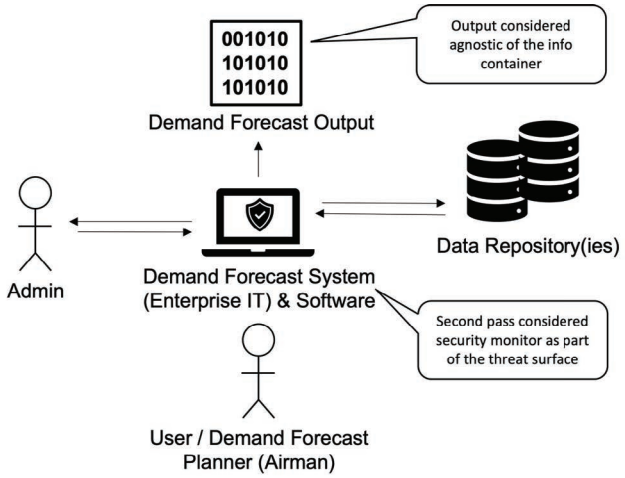
OCTAVE Risk and Mitigation Analysis Details

This appendix supplies additional details related to the risk and mitigation analysis undertaken for each of the threads in Chapter 4. These details are presented here to provide context for the summarized mitigation insights reported. For each thread, the overall scenario and diagram is reproduced here for ease of reference.

Thread 1: Demand Forecasting

Thread 1 draws on current operations by examining the use of a system (demand forecast system) by a user or demand forecast planner (airman) employing software to develop a demand forecast output. This is accomplished by accessing one or more data repositories. Additionally, it is assumed that the system is administered by one or more admin users. This is depicted in Figure E.1.

Figure E.1. Depiction of the Misuse Cases Identified as Part of Thread 1, Step 4



Risk Analysis

For this risk analysis, attention was placed on Pool 1 and Pool 2 risks, which constitute the upper third of the risk pool (4 and 31 risks, respectively). Because of the size of Pool 2, this was further subdivided into Pool 2 Upper and Pool 2 Lower risks, respectively.

- Pool 1 (4 risks) focused on the demand forecast software, which scored the highest in potential impact to mission in the Step 7 analysis. This can largely be attributed to the team's collective feeling that such errors are difficult to identify, especially in COTS software where the government often does not have visibility into the software or its development process. In particular, Pool 1 risks focused on the alteration of this software via supply-side or man-in-the-middle means.[168]
- Pool 2/Upper (14 risks) continued the focus on the demand forecast software, with more focus on potential vulnerabilities. In addition, risks relative to the ability to generate or alter the output of the demand forecast software were prevalent in Pool 2, highlighting the potential for alteration or forgery. After the demand forecast software, the demand forecast itself ranked high relative to potential mission impact because of the potential for widespread disruption.
- Pool 2/Lower (17 risks) continued the theme of risks to the demand forecast, considering a broader array of vectors that includes misuse of credentials, known and unknown vulnerabilities, and the potential for a compromised security monitor to act as a vector for attack.

Before discussing potential mitigations and our analysis of these risks, it is useful to examine some of the findings of this analysis.

- Notably, risks related to input data did not feature in the upper risk pools. This is the only information asset that does not feature in Pool 1 or 2. As part of our consequence analysis in Step 7, these items were evaluated as having a low consequence, in part because of their ever-changing and distributed nature: Their scope is generally limited (by part or platform) and their values subject to various existing checks and balances because of the variable quality and timeliness of these data. These challenges were identified in prior RAND research and echoed by domain experts.[169]
- The introduction of a security monitor as an information asset resulted in Pool 2 risks not present in the first pass of the analysis. This highlights the need to consider security controls as part of the threat surface, as evidenced by recent intrusions that have made use of endpoint detection and monitoring tools.

Mitigation Analysis

We considered 20 different controls with the potential to address the risks identified. These controls were not intended to be comprehensive but instead represent common information security and security engineering controls that affect system integrity. Therefore, they do not serve as an

[168] The alteration of software updates and the management of vulnerabilities within software libraries have been at the heart of two of the largest breaches to affect DoD: The SolarWinds and Log4J events have placed focus on the software supply chain as a vector for attack. While in neither case was information made public that indicated integrity attacks occurred as a result of these intrusions, the access achieved by the threat actors would have enabled such actions and rendered them very difficult, and perhaps even impossible, to detect.

[169] Snyder et al., 2015.

exhaustive analysis of alternatives, but rather specific options selected by SMEs that address identified concerns. As this portion of the effort was conducted prior to the receipt of the system security plan (SSP) documentation for the ESCAPE system,[170] the identified controls additionally provided a natural basis of comparison between existing controls and integrity-based threats.

Starting with the comparison with the existing cyber posture, there were controls identified in our analysis that were already partially or totally inherited by ESCAPE (approximately 10 out of 23) and several additional controls not on our list that were in place (primarily addressing confidentiality and integrity concerns not considered here).[171] Of the controls identified in our analysis that were not implemented, the items identified with the most potential for impact include (items related to NIST SP 800-53 rev. 5 controls are shown in brackets) the following:

- *Requiring bills of materials from vendors [SC-16]*. SBOMs are a current research topic generating a lot of interest for the promise of providing a computable method of identifying components that might be problematic or in need of update.[172] Related to this is the levy of requirements on the vendor to maintain a level of security guarantee [SA-4].
- *Requiring and verifying all code received from the vendor for delivery, installation, and update [SA-10(1), SA-10(6), MA-3]*. This refers to the practice of using hashes and cryptographic signatures as part of a process to validate software as supplied from the vendor and across installations. It is, however, limited to the assurance that can be provided by the generator or approver of the software.[173]
- *Implementing start-up and runtime attestation of software [IA 3(1)]*. One means of ensuring software is unchanged from compile time is to use such cryptographic verification as provided by Trusted Platform Modules. However, such an approach could incur additional monetary costs, affect performance and maintainability, and be difficult to implement in some environments.
- *Conducting analysis for vulnerabilities*. This could be done by the vendor, the government, or (ideally) both, through several mechanisms:
  - *Static Analysis [SA-11(1)]* on source and raw artifacts. This is generally accomplished using tools, although manual processes that explore both code and artifacts (such as architecture analysis) can provide greater assurance (although often at a greater cost and

[170] "Enterprise Supply Chain Analysis, Planning and Execution—Operational Data Store (ESCAPE-ODS) System Security Plan (SSP)," Version 1.2.1, October 5, 2021, Not available to the general public.

[171] This number is approximate because the control categories identified in SP 800-53 rev. 5 contain an amount of overlap and may be implemented, tailored, or inherited. The number might be higher, but the purpose of the comparison was to highlight areas that might not be considered or might not be implemented with the goal of integrity rather than to assess current controls or posture.

[172] The Cybersecurity and Infrastructure Agency has an SBOM initiative to promote this practice ("Software Bill of Materials," undated).

[173] As we were finalizing this report, the Office of Management and Budget issued memorandum M-22-18, titled *Enhancing the Security of the Software Supply Chain Through Secure Software Development Practices*. This memorandum provides guidance directing each federal agency to comply with Executive Order 14028 through two specific actions, summarized as "obtain self-attestation before deployment" and "obtain software artifacts to demonstrate conformance" (with SBOMs provided as an example)—corresponding to bullets 1 and 2 of this section (Office of Management and Budget, *Enhancing the Security of the Software Supply Chain Through Secure Software Development Practices*, M-22-18, September 14, 2022).

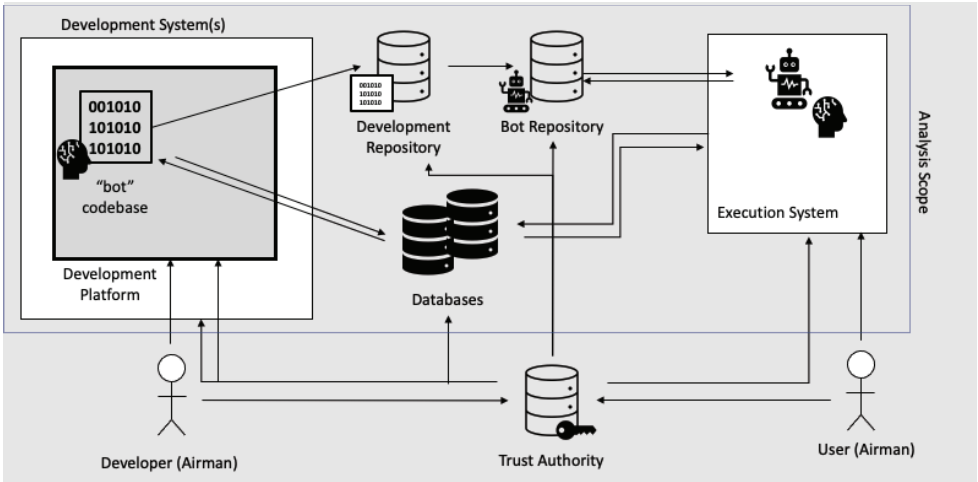
level of expertise). Mitigation capability depends on the ability to act on results but requires access to these materials.
  - *Dynamic analysis [SA-11(8)]*, which uses compiled code and so can be accomplished independently of the vendor (potentially as part of a verification process, as described above) and requires only the delivered executable.
  - *Red-teaming/penetration testing [CA-8]*, ideally using external parties.

- *Placing analysis out of band [CA-7, PM-31]*. Recognizing that monitors pose a risk because of their level of system execution and expansive data access, placing these items out of band (e.g., behind a one-way diode) from mission software will limit their ability to serve as vectors of attack. This, however, comes with costs and could further limit the ability to react to events in a timely manner.
- *Cryptographic protection of software outputs [SC-13]*. Although this control is identified as inherited (presumably because of the encryption at rest supplied by AFNet machines), encryption of individual outputs would further limit the damage a malign actor (such as an insider) might enact. Such a change would presumably require updates to the demand forecast software itself, through either HAF/A4P or vendor implementation.
- *2FA [IA-2]*. This is, again, an inherited control; however, the SSP leaves open the possibility that while the machine running the demand forecast software may use 2FA (in the form of CACs), it is likely that the demand forecast software itself does not. This opens both insider and remote threats, potentially allowing anyone with AFNet access to this software.

Thread 2: Bot Development and Employment

Thread 2 describes a robust bot development and use environment, as depicted in Figure E.2.

Figure E.2. Notional Bot Development and Employment Scenario Used for Analysis



Risk Analysis

The result of the misuse analysis for this thread was the identification of 70 total risks. As before, these risks were scored and pooled, resulting in 3 risks in Pool 1 and 28 risks in Pool 2 (again separated into Pool 2/Upper and Pool 2/Lower, with 10 and 18, risks respectively).

- Pool 1 (3 risks) focused on the bot code, which was rated in Step 7 as the information asset with the potential for severe integrity-based disruption. The threats identified in this pool related specifically to the targeting of bot code in development, by either internal or external actors (with the former incorporating both malicious insiders and threats without malicious intent, occurring through human error or inattention).
- Pool 2/Upper (10 risks) includes additional risks to the bot code, as well as risks related to the compiled bot and the data. The common theme in this risk pool is the alteration or substitution of information during testing or operations. Threats to the code while in transit are also identified here.
- Pool 2/Lower (18 risks) rounds out the focus on the compiled bot and data, with risks focused on threats to related to storage in the bot repository for the former and the falsification of data for the latter.

It is again notable that one information asset, the bot model, does not appear in either of the top risk pools. This is likely because of the stated assumption of a static model, rendering risks to the model secondary to risks to the overall code. Even without this stipulation, although attacks on learning-based models are a popular topic, their implementation in practice remains a challenging proposition.

Mitigation Analysis

Using a pool of roughly 20 controls similar to Threads 1 and 2 as a basis, we employed sources of low- and no-code security advice to guide the list of investigated controls—notably, "OWASP Top 10 LCNC Security Risks." The risks identified by OWASP include a mixture of technical, training, and procedural activities. These were further augmented by the research team's experience (OWASP and SP 800-53 rev. 5 linkages are shown in brackets).

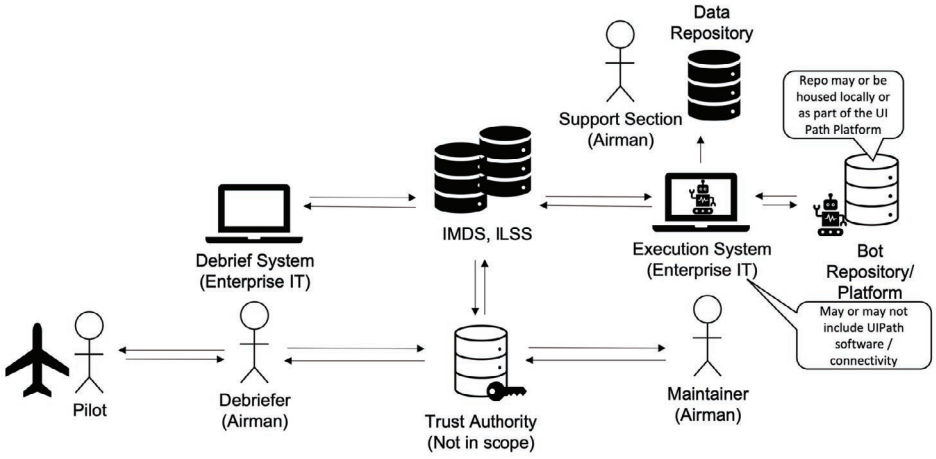
- As in prior threads, audits *[OWASP LCNC-SEC-01/10, AU-2, AU-14]* provided the widest impact, although not the largest reduction. Our analysis also recognized the *development of policy and guidance [SR-1, SA-1]* as an ongoing DAF initiative in this space and one that can be mapped to a wide variety of risks.
- Along with the need for standards comes the need to develop and enforce secure coding/best-practice guidance and training of staff *[OWASP LCNC-SEC-02/04/05/06, SA-16]*. Although the promise of low- and no-code environments is their utility to nonprogrammers, it will still be essential to maintain a level of quality in bot development operation to limit the introduction of vulnerabilities or errors leading to compromises of integrity.
- Some of the largest impacts, both per risk and across risks, focused on security of the development environment.

  - Implementing and executing a patching program for the bot platform and associated repository [MA-3(6)] addressed some of the highest risks. Although the specific configuration will drive the exact risk posture in this case, the existence of these vulnerabilities presents a point of increased attention for DAF adoption.
  - Implementing 2FA for development and storage [IA-2]. Once again, specific implementation details will drive the nature of this risk. Controlling access to the bot development and storage platforms will be a key, but potentially difficult, endeavor—especially for development platforms not fully on premises or within the DoD cloud.
  - Conducting reviews on bot products to ensure least privilege (such as through code review [OWASP LCNC-SEC-01/02, AC-6]) and executing red-team assessments of both the development and operational environment [CA-8].
- Relatively inexpensive controls related to bot design and development recommended by OWASP include ensuring that user and service accounts are decoupled [OWASP LCNC-SEC-01, AC-3, AC-5] and shared connections are limited and monitored [OWASP LCNC-SEC-02/07, CA-3].

Thread 3: Bot Employment for Data Integration to Enable Failure Analysis

Thread 3 brings together concepts from Threads 1 and 2 to examine the scenario outlined in Chapter 3, as depicted in Figure E.3.

Figure E.3. Depiction of the Thread 3 Bot Use Case



Risk Analysis

Given the relationship of this thread to the prior threads, we were able to map the information assets identified to those already examined in prior threads (Table E.1), allowing us to reuse the threat and risk analysis already developed.

Table E.1. Mapping of Information Assets and Areas of Concern for Thread 3, Based on Thread 1 and Thread 2 Analysis

Info Assets	Mapped to	Misuse Cases
Bot (compiled)	Thread 2: Bot (compiled) Thread 1: Demand forecast software	Alteration of the bot executable via injection, alteration, deletion to undermine execution (mission environment)
		Alteration of the bot executable via injection, alteration, deletion during submission of the bot to the repository
		Substitution of artifacts/data in transit
Data (operational)	Thread 2: Data (operational)	Alteration/faking of operational data (at rest)
		Substitution of (altering, faking) artifacts/data in transit
Bot output	Thread 1: Demand forecast output	Generation of a fake (output) by executing the (bot) using incorrect/spurious data
		Generation of a fake (output) by altering an existing [output]

Unlike the other threads that focused broadly on supply chain functions, the nature of the scenario and the HAF/A4 movement toward the UiPath product allowed a more targeted analysis in Step 3 (Table E.2), providing a richer context for recommendations that informed risks related to development and deployment scenarios. Using both the information asset and container mappings alongside the prior threat and risk analysis allowed the research team to quickly identify relevant risks for further consideration.

Table E.2. Information Container Mappings (with Deployment Options Highlighted)

Container Description	Owner(s)
Information Asset Risk Environment Map (Technical) 9a	
Internal	
Debrief system (IT)	AFNet
Execution system (IT)	AFNet
IMDS*	AFNet/Cloud
ILS-S*	AFNet/Cloud
Data repository*	AFNet/Cloud

Container Description	Owner(s)
Bot platform/repository*	AFNet/UiPath
External	
Bot platform/repository*	UiPath
Information Asset Risk Environment Map (Physical) 9b	
Internal	
Maintenance site	USAF
External	
Cloud	USAF/Commercial
Information Asset Risk Environment Map (People) 9c	
Internal	
Pilot	USAF
Debriefer	USAF
Maintainer	USAF
Support section	USAF
External	
UiPath development and support	UiPath

NOTE: * Denotes that the item may be internal or external based on architecture or configuration.

Using the same analysis and pooling method applied in the prior analysis, approximately 50 risks from Threads 1 and 2 were identified as being relevant:

- Pool 1 (5 risks) related to supply attacks on the bot and software platform.
- Pool 2/Upper (13 risks) primarily focused on the use of known vulnerabilities to undermine the bot and bot output.
- Pool 2/Lower (18 risks) represented the largest set, primarily focused on the data and bot output information assets.
- The remaining risks were allocated into Pools 3 and 4.

Unlike Thread 1 and Thread 2, the distribution of risks in the analysis for Thread 3 spanned each information asset and threat vector within the upper risk pools while capturing many of the same issues raised in those threads.

Risk Analysis

With the mitigations already examined in Threads 1 and 2, their analysis in the context of Thread 3 is fully described by the observations detailed in Chapter 4.

Abbreviations

2FA	two-factor authentication
A4	logistics, engineering, and force protection
A4L	Logistics Directorate
AFLCMC	Air Force Life Cycle Management Center
AFMC	Air Force Materiel Command
AFNet	Air Force Network
AFRL	Air Force Research Laboratory
AFSC	Air Force Sustainment Center
AI	artificial intelligence
APSR	accountable property system of record
ATO	authority to operate
BOSS	Base Operating Stock Specialist
C3I	command, control, communication, and intelligence
CAC	common access card
CEMS	comprehensive (or centralized) engine management system
CIMIP	Comprehensive Inventory Management Improvement Plan
CIO	Chief Information Officer
COE	Center of Excellence
COTS	commercial off the shelf
COVID-19	coronavirus disease 2019
DAF	Department of the Air Force
DFA	demand forecast accuracy
DLA	Defense Logistics Agency
DLM	depot-level maintenance
DLR	depot-level reparable
DoD	U.S. Department of Defense
ECSS	Expeditionary Combat Support System
EOH	engine overhaul
ERP	enterprise resource planning system
ESCAPE	Enterprise Supply Chain Analysis, Planning, and Execution
ETM	Electronic Technical Manual
EW	electronic warfare
FY	fiscal year
GAO	Government Accountability Office

GB	Business Enterprise Systems Directorate
HAF/A4	Headquarters U.S. Air Force Deputy Chief of Staff for Logistics, Engineering and Force Protection
HN	Command, Control, Communications, Intelligence and Networks Directorate
IA	intelligent automation
IBM	International Business Machines
ILS-S	integrated logistics system supply
IMDS	integrated maintenance data system
IPB	illustrated parts breakdown
IT	information technology
JCN	job control number
LCNC	Low-Code/No-Code
LIMS-EV	Logistics, Installations, and Mission Support—Enterprise View
LMI	Logistics Management Institute
MICAP	mission impaired capability awaiting parts
ML	machine learning
MMH	maintenance man-hours
NDAA	National Defense Authorization Act
NHA MISTR	next higher assembly management of items subject to repair
NIIN	national item identification number
NIST	National Institute of Standards and Technology
NLP	natural language processing
NMCS	not mission capable for supply
NN	neural network
NSN	National Stock Number
OCTAVE	Operationally Critical Threat, Asset, and Vulnerability Evaluation
OEM	original equipment manufacturer
OIM	organization and intermediate maintenance
OSD	Office of the Secretary of Defense
OWASP	Open Web Application Security Project
PAF	Project AIR FORCE
PDM	programmed depot maintenance
PKI	public key infrastructure
PNG	Peak Policy and Next Gen
PTC	Parametric Technology Corporation
RMF	risk management framework
RMS	Requirements Management System
ROI	return on investment
RPA	robotic process automation
RPAaaS	robotic process automation as a service

SAF/FM	Secretary of the Air Force Financial Management & Comptroller
SBOM	software bill of materials
SCMS	supply chain management squadron
SCMW	supply chain management wing
SCOS	supply chain operations squadron
SEM	scheduled event maintenance
SME	subject-matter expert
SoS	source of supply
SP	Special Publication
SPM	Service Parts Management
SSDLC	system security development life cycle
SSP	system security plan
USAF	U.S. Air Force
WUC	work unit code

References

"448th Supply Chain Management Wing," webpage, Tinker Air Force Base, undated. As of November 20, 2023: https://www.tinker.af.mil/Units/AFSC/448th-Supply-Chain-Management-Wing/

"2022 Gartner Magic Quadrant for Robotic Process Automation," webpage, UiPath, undated. As of September 28, 2022: https://www.UiPath.com/resources/automation-analyst-reports/gartner-magic-quadrant-robotic-process-automation

Abel, Tony, and Ben Franjesevic, "Who Is Watching the Bots? Part 2: Operational Challenges (and Solutions)," Protiviti, June 11, 2019.

Abhimanyu V., "What Are the Limitations of RPA?" Tutorialspoint, December 8, 2022.

Accenture Federal Services, *Sustainment Lifecycle Phase Forecasting and the Impact on Business Outcomes*, July 2013, Not available to the general public.

AFMC Manual—*See* Air Force Materiel Command Manual.

Air Force Financial Systems Operations, *AFFSO Automation Development Lifecycle*, PowerPoint presentation, February 12, 2020, Not available to the general public.

Air Force Materiel Command Manual 23-101 Volume 1, *Materiel Management General D200A/N Information*, Department of the Air Force, November 17, 2016.

Air Force Materiel Command Manual 23-101 Volume 5, *Equipment Specialist Data and Reports (D200A, D200N)*, Department of the Air Force, December 15, 2021.

Alberts, Christopher, Peter Gordon, and Audrey Dorofee, *Managing Information Security Risks: The OCTAVE (SM) Approach*, Addison-Wesley Professional, 2002.

Amirkolaii, K. Nemati, A. Baboli, M. K. Shahzad, and R. Tonadre, "Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by Using Artificial Intelligence (AI)," *International Federation of Automatic Control-PapersOnLine*, Vol. 50, No. 1, July 2017.

Atchley, Walter D., Dorothy M. Clark, Salvatore J. Culosi, Lori Dunch, Robert C. Kline, Thomas E. Lang, Randy L. Moller, Matthew R. Peterson, and Michael R. Pouy, *Lifecycle Forecasting Improvement: Causative Research and Item Introduction Phase*, Logistics Management Institute, Report DL920T1, November 2010.

Babai, M. Z., A. Tsadiras, and C. Papadopoulos, "On the Empirical Performance of Some New Neural Network Methods for Forecasting Intermittent Demand," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 31, No. 3, July 2020.

Bacchetti, Andrea, and Nicola Saccani, "Spare Parts Classification and Demand Forecasting for Stock Control: Investigating the Gap Between Research and Practice," *Omega*, Vol. 40, No. 6, December 2012.

Bachman, Tovey C., Pamela J. Williams, Kristen M. Cheman, Jeffrey Curtis, and Robert Carroll, "PNG: Effective Inventory Control for Items with Highly Variable Demand," *Interfaces*, Vol. 46, No. 1, 2015.

Baisariyev, M., A. Bakytzhanuly, Y. Serik, B. Mukhanova, M. Z. Babai, M. Tsakalerou, and C. T. Papadopoulos, "Demand Forecasting Methods for Spare Parts Logistics for Aviation: A Real-World Implementation of the Bootstrap Method," *Procedia Manufacturing*, Vol. 55, 2021.

Blumberg Advisory Group, *Spare Parts Management Software State of the Art Benchmark Evaluation*, 2020.

Bornet, Pascal, "A Framework for Explaining the Power of Intelligent Automation," *Wevolver*, February 15, 2022.

"Bot," Merriam-Webster, webpage, undated. As of September 28, 2022: https://www.merriam-webster.com/dictionary/bot

Bower, Anthony G., and Steve Garber, *Statistical Forecasting of Bankruptcy of Defense Contractors: Problems and Prospects*, RAND Corporation, MR-410-AF, 1994. As of November 16, 2023: https://www.rand.org/pubs/monograph_reports/MR410.html

Boylan, John E., and Aris A. Syntetos, "Spare Parts Management: A Review of Forecasting Research and Extensions," *International Journal of Imaging Systems and Technology Journal of Management Mathematics*, Vol. 21, No. 3, November 12, 2009.

Brown, B. B., and M. A. Geisler, *Analysis of the Demand Patterns for B-47 Airframe Parts at Air Base Level*, RAND Corporation, RM-1297, 1954. As of November 16, 2023: https://www.rand.org/pubs/research_memoranda/RM1297.html

Brown, Bernice B., *Characteristics of Demand for Aircraft Spare Parts*, RAND Corporation, R-292, 1956. As of November 16, 2023: https://www.rand.org/pubs/reports/R292.html

Butler, Dwayne M., Anthony Atler, Stephen M. Worman, Lily Geyer, and Bonnie Magnuson, *Identifying Efficiencies in the Supply Chain for Training Ammunition: Methods, Models, and Recommendations*, RAND Corporation, RR-952-A, 2016. As of November 16, 2023: https://www.rand.org/pubs/research_reports/RR952.html

Caralli, Richard A., James F. Stevens, Lisa R. Young, and William R. Wilson, *Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process*, Software Engineering Institute, Carnegie Mellon University, May 2007.

Centers for Disease Control and Prevention, "CDC Museum COVID-19 Timeline," webpage, last reviewed March 15, 2023. As of November 16, 2023: https://www.cdc.gov/museum/timeline/covid19.html

Chenoweth, Mary E., Jeremy Arkes, and Nancy Y. Moore, *Best Practices in Developing Proactive Supply Strategies for Air Force Low-Demand Service Parts*, RAND Corporation, MG-858-AF, 2010. As of November 16, 2023: https://www.rand.org/pubs/monographs/MG858.html

Choi, Boram, and Jong Hwan Suh, "Forecasting Spare Parts Demand of Military Aircraft: Comparisons of Data Mining Techniques and Managerial Features from the Case of South Korea," *Sustainability*, Vol. 12, No. 15, July 2020.

Chu, Fox, Sven Gailus, Lisa Liu, and Liumin Ni, "The Future of Automated Ports," McKinsey & Company, December 4, 2018.

Cloud One, homepage, U.S. Department of Defense, undated. As of September 28, 2022: https://cloudone.af.mil/

Croston, J. D., "Forecasting and Stock Control for Intermittent Demands," *Operational Research Quarterly*, Vol. 23, No. 3, September 1972.

Cybersecurity and Infrastructure Security Agency, "Software Bill of Materials (SBOM)," webpage, undated. As of November 16, 2023: https://www.cisa.gov/sbom

Dastin, Jeffrey, "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women," Reuters, October 10, 2018.

De Gooijer, Jan G., and Rob J. Hyndman, "25 Years of Time Series Forecasting," *International Journal of Forecasting*, Vol. 22, No. 3, 2006.

Deb Roy, Suman, "What Bots May Come: An In Depth Discussion of a Learning Architecture for the Next Paradigm," *Chatbots Magazine*, March 20, 2016.

Defense Logistics Agency, "National Stock Numbers (NSNs)," webpage, undated. As of February 21, 2024: https://www.dla.mil/Disposition-Services/DDSR/Quick-Links/NSNs

DeFrank, Joshua D., "A Condition Based Maintenance Approach to Forecasting B-1 Aircraft Parts," Air Force Institute of Technology, March 3, 2017.

Department of Defense Instruction 8510.01, *Risk Management Framework (RMF) for DoD Information Technology (IT)*, U.S. Department of Defense Office of the Chief Information Officer, July 19, 2022.

Department of Defense Manual 4140.01 Volume 2, *DoD Supply Chain Materiel Management Procedures: Demand and Supply Planning*, U.S. Department of Defense, 2018.

Department of the Air Force Guidance Memorandum 2021-01 to Department of the Air Force Manual 23-122, *Materiel Management Procedures*, July 7, 2021.

DoD—*See* U.S. Department of Defense.

Donaldson, Michele, "Team Illuminates Supply Risks That Impact Defense," Eglin Air Force Base, August 31, 2022.

Duane, J. T., "Learning Curve Approach to Reliability Monitoring," *Institute of Electrical and Electronic Engineers Transactions on Aerospace*, Vol. 2, No. 2, April 1964.

"Enterprise Supply Chain Analysis, Planning and Execution—Operational Data Store (ESCAPE-ODS) System Security Plan (SSP)," Version 1.2.1, October 5, 2021, Not available to the general public.

Feeney, G. J., and C. C. Sherbrooke, *Systems Analysis and Supply Management*, RAND Corporation, RM-4054-PR, 1964. As of November 16, 2023: https://www.rand.org/pubs/research_memoranda/RM4054.html

GAO—*See* Government Accountability Office.

Gartner, "Gartner Magic Quadrant and Critical Capabilities: Methodologies Evolution," September 10, 2019.

Government Accountability Office, *Defense Inventory: Opportunities Exist to Save Billons by Reducing Air Force's Unneeded Spare Parts Inventory*, GAO-07-232, April 27, 2007.

Government Accountability Office, *Defense Inventory: Management Actions Needed to Improve the Cost Efficiency of Navy's Spare Parts Inventory*, GAO-09-103, December 12, 2008.

Government Accountability Office, *Defense Inventory: Army Needs to Evaluate Impact of Recent Actions to Improve Demand Forecasts for Spare Parts*, GAO-09-199, January 12, 2009.

Government Accountability Office, *Defense Inventory: Defense Logistics Agency Needs to Expand on Efforts to More Effectively Manage Spare Parts*, GAO-10-469, May 11, 2010.

Government Accountability Office, *DOD's 2010 Comprehensive Inventory Management Improvement Plan Addressed Statutory Requirements, but Faces Implementation Challenges*, GAO-11-240R, January 7, 2011.

Government Accountability Office, *High Risk Series: An Update*, GAO-15-290, February 11, 2015.

Government Accountability Office, *Progress on Many High-Risk Areas, While Substantial Efforts Needed on Others*, GAO-17-317, February 15, 2017.

Gružauskas, Valentas, and Diwakaran Ragavan, "Robotic Process Automation for Document Processing: A Case Study of a Logistics Service Provider," *Journal of Management*, Vol. 36, No. 2, December 2020.

Guo, Feng, Jun Diao, Qiuhong Zhao, Dexin Wang, and Qiang Sun, "A Double-Level Combination Approach for Demand Forecasting of Repairable Airplane Spare Parts Based on Turnover Data," *Computers & Industrial Engineering*, Vol. 110, August 2017.

HAF/A4 Logistics Automation, "Technology Modernization Fund: Full Project Proposal," PowerPoint presentation, October 2021.

Heitzenrater, Chad, and Andrew Simpson, "Misuse, Abuse and Reuse: Economic Utility Functions for Characterising Security Requirements," in *Proceedings of the 2nd International Workshop on Agile Secure Development*, August 2016.

Hodges, James S., and Raymond A. Pyles, *Onward Through the Fog: Uncertainty and Management Adaptation in Systems Analysis and Design*, RAND Corporation, R-3760-AF/A/OSD, 1990. As of November 17, 2023:
https://www.rand.org/pubs/reports/R3760.html

Hyndman, Rob J., "Another Look at Forecast-Accuracy Metrics for Intermittent Demand," *Foresight*, No. 4, June 2006.

IBM—*See* International Business Machines.

Interagency Task Force in Fulfillment of Executive Order 13806, *Assessing and Strengthening the Manufacturing and Defense Industrial Base and Supply Chain Resiliency of the United States*, September 2018.

International Business Machines, *Cost of a Data Breach Report 2022*, 2022.

Jabbour, Kamal, and Sarah Muccio, "The Science of Mission Assurance," *Journal of Strategic Security*, Vol. 4, No. 2, 2011.

Kordy, Barbara, Ludovic Piètre-Cambacédès, and Patrick Schweitzer, "DAG-Based Attack and Defense Modeling: Don't Miss the Forest for the Attack Trees," *Computer Science Review*, Vols. 13–14, November 2014.

Kramer, Jeff, "Is Abstraction the Key to Computing?" *Communications of the Association for Computing Machinery*, Vol. 50, No. 4, April 2007.

Lebeuf, Carlene, *A Taxonomy of Software Bots: Towards a Deeper Understanding of Software Bot Characteristics*, master's thesis, University of Victoria, 2018.

Lebeuf, Carlene, Alexey Zagalsky, Matthieu Foucault, and Margaret-Anne Storey, "Defining and Classifying Software Bots: A Faceted Taxonomy," *2019 Institute of Electrical and Electronic Engineers/Association for Computing Machinery 1st International Workshop on Bots in Software Engineering (BotSE)*, 2019.

Lee, Hanjun, and Jaedong Kim, "A Predictive Model for Forecasting Spare Parts Demand in Military Logistics," *2018 Institute of Electrical and Electronic Engineers International Conference on Industrial Engineering and Engineering Management (IEEM)*, December 2018.

Leonard, Andrew, *Bots: The Origin of New Species*, Penguin Books Limited, 1998.

Light, Thomas, Michael Boito, Tim Conley, Larry Klapper, and John Wallace, *Understanding Changes in U.S. Air Force Aircraft Depot-Level Reparable Costs over Time*, RAND Corporation, 2018, Not available to the general public.

Loredo, Elvira N., John F. Raffensperger, and Nancy Y. Moore, *Measuring and Managing Army Supply Chain Risks: A Quantitative Approach by Item Number and Commercial Entity Code*, RAND Corporation, RR-902-A, 2015. As of November 17, 2023: https://www.rand.org/pubs/research_reports/RR902.html

Makridakis, Spyros, and Michèle Hibon, "The M3-Competition: Results, Conclusions and Implications," *International Journal of Forecasting*, Vol. 16, No. 4, 2000.

"Maritime Port Operators See Great Promise in Artificial Intelligence," *Supply Chain Quarterly*, September 20, 2019.

McClausland, Tammy, "The Bad Data Problem," *Research-Technology Management*, Vol. 64, No. 1, 2021.

"Measuring RPA ROI—How to Do It Right?" *Digital Workforce* blog, July 1, 2020. As of September 27, 2022: https://digitalworkforce.com/rpa-news/measuring-rpa-roi-how-to-do-it-right/

Mehta, Aaron, "How Coronavirus Could Impact the Defense Supply Chain," Defense News, March 20, 2020.

Miller, Amanda, "Half of Air Force Advanced STEM Billets Go Unfilled or Require Waivers," *Air Force Magazine*, August 21, 2022.

Moore, Nancy Y., Elvira N. Loredo, Amy G. Cox, and Clifford A. Grammich, *Identifying and Managing Acquisition and Sustainment Supply Chain Risks*, RAND Corporation, RR-549-AF, 2015. As of November 17, 2023: https://www.rand.org/pubs/research_reports/RR549.html

"Multi-Source Exploitation Assistant for the Digital Enterprise (MEADE)," webpage, SAM.gov, last updated September 30, 2022. As of November 17, 2023: https://sam.gov/opp/84bce1a9a5ba460a9e6530f7ff70eb38/view

National Institute of Standards and Technology, *Security and Privacy Controls for Information Systems and Organizations*, Special Publication 800-53, revision 5, September 2020.

National Vulnerability Database, National Institute of Standards and Technology, undated. As of November 17, 2023:
https://nvd.nist.gov/vuln

Naz, Farheen, Anil Kumar, Abhijit Majumdar, and Rohit Agrawal, "Is Artificial Intelligence an Enabler of Supply Chain Resiliency Post COVID-19? An Exploratory State-of-the-Art Review for Future Research," *Operations Management Research*, Vol. 15, Nos. 1–2, September 2021.

NIST—*See* National Institute of Standards and Technology.

Nystrom, Bill, *Democratizing Automation for Every Airman*, PowerPoint presentation, UiPath, undated.

O'Connell, Caolionn, Bryan Boling, Jonathan Balk, James R. Broyles, and Monika Cooper, *Hidden Disruptions to the Supply Chain: Resistance Is Futile*, RAND Corporation, 2021, Not available to the general public.

O'Connell, Caolionn, Elizabeth Hastings Roer, Rick Eden, Spencer Pfeifer, Yuliya Shokh, Lauren A. Mayer, Jake McKeon, Jared Mondschein, Phillip Carter, Victoria A. Greenfield, and Mark Ashby, *Managing Risk in Global Supply Chains*, RAND Corporation, RR-A425-1, 2021. As of November 16, 2023:
https://www.rand.org/pubs/research_reports/RRA425-1.html

Office of Management and Budget, *Enhancing the Security of the Software Supply Chain Through Secure Software Development Practices*, M-22-18, September 14, 2022.

"Ombuds: Guiding Principles," webpage, Gartner, undated. As of September 28, 2022:
https://www.gartner.com/en/about/ombuds/ombuds-guide

O'Neal, Thomas R., *Sortie-Based Aircraft Component Demand Rate to Predict Requirements*, Air Force Institute of Technology, March 2020.

Open Worldwide Application Security Project, "OWASP Top 10 Low-Code/No-Code Security Risks," webpage, undated. As of August 24, 2022:
https://owasp.org/www-project-top-10-low-code-no-code-security-risks/

OWASP—*See* Open Worldwide Application Security Project.

Parametric Technology Corporation, "Case Studies," webpage, undated. As of December 6, 2023:
https://www.ptc.com/en/case-studies

Parikh, Nish, "Understanding Bias in AI-Enabled Hiring," *Forbes*, October 14, 2021.

Peltz, Eric, Amy G. Cox, Edward W. Chan, George E. Hart, Daniel Sommerhauser, Caitlin Hawkins, and Kathryn Connor, *Improving DLA Supply Chain Agility: Lead Times, Order Quantities, and Information Flow*, RAND Corporation, RR-822-OSD, 2015. As of November 16, 2023:
https://www.rand.org/pubs/research_reports/RR822.html

Pinçe, Çerağ, Laura Turrini, and Joern Meissner, "Intermittent Demand Forecasting for Spare Parts: A Critical Review," *Omega*, Vol. 105, No. 1, July 2021.

Posadas, Sergio, Carl M. Kruger, Catherine M. Beazley, Russell S. Salley, John A. Stephenson, Esther C. Thron, and Justin D. Ward, "Forecasting Parts Demand Using Service Data and Machine Learning," Logistics Management Institute, January 2020.

"Positioning Technology Players Within a Specific Market," webpage, Gartner, undated. As of September 28, 2022:
https://www.gartner.com/en/research/methodologies/magic-quadrants-research

Prather, Kayla, "AFMC Robotic Process Automation Roadshow Drives Innovation," Air Force Materiel Command, December 15, 2021.

Public Law 111-84, National Defense Authorization Act for Fiscal Year 2010, October 28, 2009.

RPA Center of Excellence, "RPA Center of Excellence (COE) Accomplishments," PowerPoint presentation, March 2022.

Sangster, Douglas, and Benjamin Young, "PartBot (for Maintainers, by Maintainers)," PowerPoint presentation, 35th Maintenance Group, undated. As of September 28, 2022: https://view.highspot.com/viewer/60d8e774c79c5250dc4ed0b0

Schneier, Bruce, "Attack Trees," *Schneier on Security* blog, December 1999.

Shanbhag, Abhishek, "Why the Future of Chatbots Is Low Code," *BotCore by Acuvate* blog, January 15, 2021.

Sherbrooke, Craig C., *Using Sorties vs. Flying Hours to Predict Aircraft Spares Demand*, Logistics Management Institute, 1997.

Shostack, Adam, "STRIDE Chart," *Microsoft Security* blog, September 11, 2007.

Shostack, Adam, *Threat Modeling: Designing for Security*, 1st ed., Wiley Publishing, 2014.

Simchi-Levi, David, William Schmidt, Yehua Wei, Peter Yun Zhang, Keith Combs, Yao Ge, Oleg Gusikhin, Michael Sanders, and Don Zhang, "Identifying Risks and Mitigating Disruptions in the Automotive Supply Chain," *Interfaces*, Vol. 45, No. 5, October 2015.

Simoes, Bruna Sofia, "Infographic: 10 Metrics You Should Be Tracking to Drive RPA Success," Blueprint, webpage, January 22, 2021. As of September 27, 2022: https://www.blueprintsys.com/blog/rpa/10-metrics-you-should-be-tracking-to-drive-rpa-success

Sindre, Guttorm, and Andreas L. Opdahl, "Eliciting Security Requirements with Misuse Cases," *Requirements Engineering*, Vol. 10, No. 1, January 2005.

Sirdeshmukh, Swapnil, Yashdeep Saran, and Ankit Tondon, "Faster Decision-Making with RPA in High-Tech Supply Chains," Infosys, February 2, 2019.

Slota, Stephen C., Kenneth R. Fleischmann, Sherri Greenberg, Nitin Verma, Brenna Cummings, Lan Li, and Chris Shenefiel, "Good Systems, Bad Data? Interpretations of AI Hype and Failures," *Proceedings of the Association for Information Science and Technology*, Vol. 58, No. 1, 2020.

Snyder, Don, George E. Hart, Kristin F. Lynch, and John G. Drew, *Ensuring U.S. Air Force Operations During Cyber Attacks Against Combat Support Systems: Guidance for Where to Focus Mitigation Efforts*, RAND Corporation, RR-620-AF, 2015. As of November 16, 2023: https://www.rand.org/pubs/research_reports/RR620.html

"Supply Chain," webpage, Merriam-Webster, undated. As of September 29, 2022: https://www.merriam-webster.com/dictionary/supply%20chain

Syntetos, Aris A., Mohamed Zied Babai, John Boylan, Stephan Kolassa, and Konstantinos Nikolopoulos, "Supply Chain Forecasting: Theory, Practice, Their Gap and the Future," *European Journal of Operational Research*, Vol. 252, No. 1, November 2015.

Syntetos, Aris A., John E. Boylan, and J. D. Croston, "On the Categorization of Demand Patterns," *Journal of the Operational Research Society*, Vol. 56, No. 5, August 25, 2004.

"Tampering," webpage, National Institute of Standards and Technology Information Technology Laboratory Computer Security Resource Center, undated. As of September 8, 2022: https://csrc.nist.gov/glossary/term/tampering

Teodorczuk, Bart, "How to Measure RPA Success? A Guide to Robotic Process Automation Metrics," *Flobotics* blog, December 23, 2021.

Tøndel, Inger Anne, Jostein Jensen, and Lillian Røstad, "Combining Misuse Cases with Attack Trees and Security Activity Models," *Proceedings of 2010 International Conference on Availability, Reliability and Security, ARES*, 2010.

Radware, "Types of Bots: An In-Depth Guide by Radware," undated. As of November 16, 2023: https://www.radware.com/cyberpedia/bot-management/types-of-bots/

USAF—*See* U.S. Air Force.

U.S. Air Force, 72nd Air Base Wing Public Affairs, *Tinker Air Force Base 80th Anniversary Units and Mission Structure*, 2022.

U.S. Air Force, 420th Supply Chain Management Squadron, *AF DFA and Bias with IMR Charts*, PowerPoint presentation, July 11, 2022.

U.S. Air Force A4P (Program Integration Directorate), *A4 Chief Information Officer Cybersecurity Risk Frame*, September 4, 2020.

U.S. Air Force Scientific Advisory Board, *Sustaining Air Force Aging Aircraft into the 21st Century*, Department of the Air Force, August 1, 2011.

U.S. Air Force, Forecast Analysis Comparison Tool, database, 2020–2021.

U.S. Department of Defense, *Supply Chain Metrics Guide*, 3rd ed., 2021.

Van der Auweraer, Sarah, and Robert N. Boute, "Forecasting Spare Part Demand Using Service Maintenance Information," *International Journal of Production Economics*, Vol. 213, July 2019.

Van der Auweraer, Sarah, Robert N. Boute, and Aris A. Syntetos, "Forecasting Spare Part Demand with Installed Base Information: A Review," *International Journal of Forecasting*, Vol. 35, No. 1, 2019.

"What Is Intelligent Automation?" International Business Machines Cloud Education, webpage, undated. As of November 17, 2023: https://www.ibm.com/cloud/learn/intelligent-automation

"What Is Robotic Process Automation?" webpage, Association for Intelligent Information Management, undated. As of September 28, 2022: https://www.aiim.org/what-is-robotic-process-automation

The White House, *Building Resilient Supply Chains, Revitalizing American Manufacturing, and Fostering Broad-Based Growth: 100-Day Reviews Under Executive Order 14017*, June 2021.

Windsor, Sandy, "Escaping Today's Supply Chain Challenges," Air Force Sustainment Center, September 27, 2017.

Wolfe, Frank, "Joint Warfighting Concept Assumes 'Contested Logistics,'" *Defense Daily*, October 6, 2020.

Zhang, Li Ang, Yusuf Ashpari, and Anthony Jacques, *Understanding the Limits of Artificial Intelligence for Warfighters: Volume 3, Predictive Maintenance*, RAND Corporation, RR-A1722-3, 2024.