RAND

Research Report

BONNIE L. TRIEZENBERG, SARAH ZABEL, RACHEL STERATORE, ADRIAN SALAS, IVAN LEPETIC, KATIE A. WILSON, NATALIA HENRIQUEZ SANCHEZ, JAMES FAN, ALEXIS LEVEDAHL, SARAH W. DENTON

Underperforming Software and Information Technology in the Department of Defense

For more information on this publication, visit **www.rand.org/t/RRA2927-1**.

Published by the RAND Corporation, Santa Monica, Calif.

Cover: U.S. Army photo.

About This Report

This report responds to direction by the U.S. Congress through the fiscal year 2023 (FY23) National Defense Authorization Act (NDAA), Section 241, "Study on Costs Associated with Underperforming Software and Information Technology." The NDAA directed the Secretary of Defense to commission an independent study of the challenges associated with the use of software and information technology (IT) in the Department of Defense (DoD), and the effects of and potential solutions to such challenges. This report provides study results to the Chief Information Officer (CIO) for DoD.

The research reported here was completed in September 2024 and underwent security review with the sponsor and the Defense Office of Prepublication and Security Review before public release.

RAND National Security Research Division

This research was sponsored by the Chief Information Officer for the U.S. Department of Defense and conducted within the Acquisition and Technology Policy Program of the RAND National Security Research Division, which operates the National Defense Research Institute (NDRI), a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense intelligence enterprise.

For more information on the RAND Acquisition and Technology Policy Program, see www.rand.org/nsrd/atp or contact the director (contact information is provided on the webpage).

Acknowledgments

We are grateful to the numerous individuals in the office of the DoD CIO, Office of the Secretary of Defense (OSD) CIO, and the services and agencies that contributed to the development of this report, responding to interview requests and sharing performance data with us. We also thank the over 6,500 men and women who responded to our survey regarding their recent experience with DoD's information technology and software applications. This research could not have been completed without their input. We also thank our RAND colleagues who provided "in-stride" review of our research and gave us a critical sounding board as we developed our approach, analysis, and conclusions. As always, we retain full responsibility for the quality of our research.

Summary

This report presents the results of an independent study regarding the performance of information technology (IT) and software-based systems in the U.S. Department of Defense (DoD) and the impact of that performance on DoD operations and mission readiness. The study was sponsored by the DoD Chief Information Officer (CIO) based on direction in the fiscal year 2023 (FY23) National Defense Authorization Act (NDAA), Section 241.

Section 241 of the FY23 NDAA directs an independent study on the challenges associated with the use of software and information technology in DoD, the effects of such challenges, and their potential solutions.

Issue

DoD workforce responses to a formal survey, other studies, and anecdotal evidence suggest that IT infrastructure and software-based systems throughout DoD are plagued by poor performance, which has potential negative impact on institutional and operational needs. These problems are believed to come from deferred investment in hardware and software in the department, excessive complexity in how user environments are managed, and poor system design and maintenance. To date, however, there has been no comprehensive effort to measure how significant these problems truly are or how they affect the DoD mission and workforce. This study provides a first look at quantifying impacts of underperforming software on department productivity, mission readiness, and morale, in order to help DoD understand its current situation and drive measurable improvement.

Approach

Our approach centers on three tasks mandated by the NDAA language: a survey to establish the baseline understanding of the extent of the problem, discussions with the military service CIOs to identify potential causes and remedies, and development of a framework for measuring future progress against goals.

Key Findings and Recommendations

Our findings are as follows:

- Service members and civilians experience a variety of technical issues in using their DoD-provided IT and software, some of which significantly affect productivity, mission readiness, and morale.
  - A conservative lower-bound estimate of the cost to DoD of lost productivity due to IT and software issues for FY23 is $2.5 billion.

  - While the average productivity loss when using a software application rated as "critical to mission" is two hours per month, one in ten users experiences more than eight hours of productivity loss per month when interacting with a single system critical to their work.
  - After adjusting for self-selection bias, we conservatively estimate that 5 percent of the DoD workforce may be strongly motivated to depart from service due to poorly performing IT and software.
- Conditions throughout the service delivery chain contribute to these issues.
- The combination of authorities, resources, and responsibilities involved makes the problems difficult to track and resolve.
- There are significant discrepancies in the perceived mission impact of user issues between the users themselves and those responsible for providing the capability or service.

With the above in mind, we offer the following recommendations:

- Improve service and reliability for outside the continental United States (OCONUS) secret-level internet protocol router network.
- Regard virtual private networks (VPNs) or follow-on technical solutions as critical infrastructure and ensure appropriate redundancy and resilience.
- Conduct periodic reviews of the standard configuration and create scaled-down configurations that provide better performance to specific user types, including minimized start-up processing for users of shared laptops and minimized background processing and improved reliability for IT used in mission-critical environments.
- Create a reliable pipeline for timely refresh of end-user devices.
- Provide mission owners and service/capability providers throughout DoD visibility into the sources, degrees, and impacts of IT issues affecting their workforce.
- Use automated collection of IT performance data to identify the bottom 10 percent of computing environments.
- Explore additional ways to identify and resolve IT and software problems as mission or capability issues, working beyond the traditional layered help-desk structure.
- Strengthen the ability of mission owners and commanders to identify and address technological problems affecting mission accomplishment.

Contents

Figures and Tables

Figures

Tables

Chapter 1

Introduction

I often sit long
waiting for laptop to run
where is the IT?

—Major Tim Lang, USMC, Okinawa

This report presents results of an independent study regarding the performance of information technology and software systems in the U.S. Department of Defense (DoD) and the impact of that performance on DoD operations and mission readiness.[1] The study was sponsored by the DoD Chief Information Officer (CIO) in accord with direction in Section 241 of the fiscal year 2023 (FY23) National Defense Authorization Act (NDAA).[2]

Section 241 of the FY23 NDAA directed an independent study of the challenges associated with the use of software and information technology (IT) in DoD, the effects of such challenges, and potential solutions. More specifically, the NDAA directed that this independent study should include (1) a survey of members of each armed service to identify the most important software and IT challenges that result in lost work hours; (2) a summary of policy and technical challenges that limit the ability of military departments to implement needed software and IT reforms, based on interviews with military department CIOs; (3) the development of a framework for assessing underperforming software and IT; and (4) the development of recommendations to address challenges identified in the survey and to improve processes through which DoD provides software and IT. The study explicitly excluded embedded software in weapon systems. The full text of Section 241 of the NDAA is provided in Appendix E.

Information Technology Infrastructure and Application Software in Department of Defense Operations and Mission Readiness

DoD increasingly depends on software and IT infrastructure to conduct its operations. Without IT infrastructure and application software, military and civilian service members are not paid, new systems and services cannot be specified or acquired, situational awareness and updated orders cannot be efficiently communicated, and modernized logistics and weapons systems become less effective.

[1] Major Lang submitted the haiku in the epigraph in response to our survey on IT infrastructure and software. We use it as the introduction and theme of our report with his permission.

[2] Public Law 117–263, James M. Inhofe National Defense Authorization Act for Fiscal Year 2023, Section 241, Study on Costs Associated with Underperforming Software and Information Technology, December 23, 2022. Appendix E provides the full text of Section 241 of the NDAA.

Problems arise when IT infrastructure and application software are poorly designed, slow, or otherwise ineffective in helping DoD personnel complete their assigned duties. While the same may be true of many corporations and other organizational entities, DoD's size, complexity, and global reach mean that software and network performance problems can be amplified to a degree seldom seen by other enterprise systems. DoD employs more than 2.8 million active-duty personnel, National Guard and reserve forces, and government civilians.[3] While these personnel reside primarily in the United States, they deploy rapidly to remote locations across the globe, with their attendant IT systems, in a time of crisis. The size and scope of these deployments has no comparison in private industry.[4] DoD reportedly has "roughly ten thousand operational systems, thousands of data centers, tens of thousands of servers, millions of computers and IT devices, and hundreds of thousands of commercial mobile devices."[5] Furthermore, the need for multiple layers of security and the division of authorities among DoD entities results in a highly fragmented IT infrastructure.[6] Poor performance of DoD's software and networks has a direct impact on DoD's mission readiness, operations, and security, yet no systematic measures of that impact currently exist.

Members of DoD have publicly complained of underperforming software and hardware.[7] Service members and civilians often have to work with out-of-date software and hardware in environments where applications and system services compete for resources and aging equipment adversely affects system resilience. Users report waiting 30 minutes or more to log in to their machines, which sometimes crash while performing simple operations. Users also complain about old equipment and find it easier to work on a personal computer than on their government-furnished equipment. This IT environment may prevent users from working efficiently or completing their work in a timely, safe, and secure manner.

Concern for the impact of poorly performing IT and software-based systems on retention, mission, and DoD's budget has prompted several studies and action by the military services and Office of the Secretary of Defense (OSD). These studies have been limited in scope to a single service or a single issue, however, and in the FY23 NDAA, Congress directed this larger study to encompass all IT and software-based systems (excluding only embedded systems) used throughout DoD.[8]

---

[3] Defense Manpower Data Center, *Number of Military and DoD Appropriated Fund Civilian Personnel*, data set, September 30, 2023. For perspective, Walmart employs approximately 2.1 million associates worldwide. Walmart, "How Many People Work at Walmart?," webpage, undated.

[4] For example, nearly 410,000 U.S. personnel are estimated to have been deployed in support of Operation Iraqi Freedom in the spring of 2003. U.S. Central Command, Assessment and Analysis Division, *Operation IRAQI FREEDOM—By the Numbers*, 30 April 2003.

[5] DoD, *DoD Digital Modernization Strategy*, June 2019, p. 7.

[6] To illustrate, each military service and some defense or field agencies operate their own IT service departments independently of one another. Systems are usually operated by program management offices, which are typically not associated with the IT service department and not subordinate to the military service CIO. The military service CIOs are empowered by law and report to their service secretary, not to the DoD CIO.

[7] A comment from a subreddit thread went viral in late 2021: thegirlisok, "Fix our computers!," archived post from the r/Military subreddit thread, undated. Many others then picked up this theme, posting under the hashtag #fixourcomputers. This included an open letter posted on X (then Twitter) and Redditt by Michael Kanaan, then direction of operations for the Air Force—MIT Artificial Intelligence Accelerator. Michael J. Kanaan, "An Open Letter, Fix Our Computers," X.com, January 2022.

[8] Public Law 117–263, Section 241, 2022.

Definitions Used in This Report

While the NDAA language does not specifically define the words *information technology* and *software*, in this report we segregate the overall computing systems that support DoD (and thus are inclusive of the NDAA's reference to IT and software) as follows:

- *Application software* refers to software designed to handle specific tasks for the user. When we use the word *application* without a modifier, we mean application software.
- *Information technology infrastructure* consists of the supporting hardware, software, communication, and information security services that a computing system requires to operate, but that can be shared by multiple computing systems for scalability.[9] It is often useful to separate the IT infrastructure that supports networks (e.g., routers, switches, cables, and network protocol software) from that of the endpoints (e.g., servers, laptops, smartphones, and operating system software) of the network.
- *Endpoint* refers to the IT infrastructure that supports the endpoints as described above. Application software may be hosted on endpoints.
- *End-user device* refers to the laptop or mobile device through which the user interacts with application software. An end-user device is one type of endpoint.
- *Network* refers to the IT infrastructure that connects two or more computers for the purpose of communicating data electronically in a near-seamless fashion.

Research Questions and Approach

Although the NDAA language is quite specific in outlining the tasks for this study, it is less explicit regarding the research questions to be answered by those tasks. Therefore, we engaged with congressional staffers to better understand the motivations for the study and formulated the following research questions:

1. How much time does it currently take for DoD personnel to access DoD IT infrastructure and applications to begin productive work?
2. Once productive work begins, how much effort is spent resolving issues or recovering effort lost when IT infrastructure or application software malfunctions?
3. How satisfied are DoD personnel with the IT infrastructure and application software supplied to them to perform their mission-essential tasks? To perform other tasks?
4. What barriers do military department and service CIOs face when taking action to improve the performance of IT infrastructure and application software within DoD?
5. How should DoD measure the state of IT infrastructure and application software over time to support congressional assessment of (a) where investment is needed and (b) the impact of prior investments?[10]

[9] Adapted from DoDI 5000.75, *Business Systems Requirements and Acquisition*, Change 2, January 2020.

[10] Note that the NDAA has many more goals for the framework, some of which may never be feasible and others of which, while feasible in the long term, would require concerted effort and change on the part of both DoD and Congress to implement. Chapter 5 provides a fuller exploration of the feasibility of the NDAA listed goals.

We first conducted background research to understand the current state of DoD IT infrastructure and application software, as well as actions the different military departments and services are taking to measure or improve their performance. We then developed and administered a survey to a representative sample of DoD personnel to answer questions 1 through 3. In parallel with developing the survey, we engaged the military department and service CIOs (as well as a former CIO and a chief technology and innovation officer) to elicit their perspectives on both the current state of IT infrastructure and software and the barriers to positive change. We used semistructured interviews that we thematically coded and analyzed to answer question 4. We also formulated a set of metrics and a framework for making or assessing the impact of investments in DoD's IT infrastructure and application software. We requested and received briefings regarding system performance data measured by the services to validate the proposed framework concept and provide insight into question 5.

Limitations of Our Approach

There are several known limitations to this study. One of the study's stated goals is to quantify the impacts of underperforming IT and software on user productivity, retention, and mission through a broad survey of service members. This methodology depends on survey respondents' self-reported recent experiences with DoD-provided IT and software. Because factors such as the time it takes to start a computer and initiate work are self-reported rather than measured, the resulting estimates of productivity loss are neither precise nor fully accurate. The resulting estimates do, however, have validity across a wider range of DoD personnel than current efforts to measure these metrics, given that those efforts reach only a portion of the total force. The nature of a broad survey also means that other than for a few applications in wide use across the department, the quantity of data points we collected on each more sparsely used application is insufficient for meaningful statistical analysis.[11] While we can make meaningful observations about applications in general categories, we can only rarely make observations about the performance of specific applications.

Another potential issue is the loose relationship between time spent actively using a computer and user productivity. A person can be productive in a job even while waiting for a computer to initialize; likewise, time spent actively using a computer is not necessarily productive.[12] Even when time required for client device start-up, communications, and server processing can be measured, there is no guarantee that the time spent in communications or system overhead results in lost productivity, or that restoring that time would increase productivity. Having conducted the survey, however, we are less concerned with this limitation. Users tell us that wait times vary widely and are rarely predictable,

[11] In response to our request that users rate the performance of their mission-critical software applications, only 36 applications (or family of applications) received more than 20 ratings, and these make up two-thirds of the total number of ratings received. The last one-third of the ratings are for applications with fewer than 20 ratings each.

[12] In fact, several researchers have found that what frustrates users and leads to inefficiency is when the pace of software cannot keep up with the pace of operations, not the absolute duration of response times. For an introduction to this line of research, we recommend Jim Dabrowski and Ethan Munson, "40 Years of Searching for the Best Computer System Response Time," *Interacting with Computers*, Vol. 23, No. 5, 2011.

and that they are spending time actively engaging with the mal-performing systems rather than performing other tasks.

The congressional language that charters the study also reflects great concern about funds spent on DoD IT and software, with elements clearly focused on identifying waste and inefficiency in IT and software development and operations. These factors cannot be measured accurately. More specifically, the NDAA asks that the assessment framework compare costs for DoD IT and software with the equivalents from the private sector, which are not exposed with any degree of reliability or accuracy. Second, the NDAA asks the assessment framework to compare actual funds spent on IT and software with planned funds for those systems, though funds obligated on contracts (a majority of the funding in this case) are actually expended throughout the period of performance of those contracts, which may be years. Even if funds expended on programs could be tracked, funds spent on unacknowledged IT services (a.k.a. shadow IT), are not earmarked in accounting systems. As we will discuss in some depth in Chapter 3, the inability to obtain a full accounting of IT infrastructure and software spending is a source of great frustration for the military department CIOs. Much of the financial data to populate a framework as requested in the NDAA do not exist in DoD or are not centrally tracked.

Organization of This Report

Chapter 2 provides background regarding the current state of DoD IT infrastructure and application software, along with actions DoD military departments and services are taking to measure and improve it. Chapter 3 discusses insights gained from our interviews with CIOs of the military departments and services. Appendix A provides the interview protocol, coding handbook, and other methodology details. Chapter 4 then discusses the survey results and offers a perspective on how poorly performing IT and software is currently affecting productivity, mission readiness, and user satisfaction (including the issue of retention). Appendix B reproduces the survey itself and gives a detailed analysis of the responses to each question. Chapter 5 presents our proposed framework along with examples of how it could be used to measure the current state of DoD IT infrastructure and application software to assess (a) where further investment is needed and (b) the impact of past investments. Appendix C provides further detail on the structure and operation of the framework. Chapter 6 summarizes our findings and recommendations. Finally, Appendix D provides a closer look into one system, Military Health System (MHS) Genesis, that received a great deal of comment in the survey and that appears to exemplify many of the struggles of fielding complex defense business mission systems. Appendix E includes the text of NDAA Section 241, which authorizes and scopes the research reported here.

Chapter 2

Background

This issue of underperforming software and IT within DoD and the impacts it has on the ability of DoD to conduct its activities are not new, but they have become an area of increased focus. In this chapter, we review recent key studies and steps DoD has taken to understand and respond to this issue.

Defense Business Board Study

A Defense Business Board study initiated to provide recommendations to improve the user experience for basic IT services in DoD revealed the prevalence of underperforming information technology.[1] The study, released in February 2023, based its findings on interviews of senior IT leaders within DoD, a survey of 20,000 Joint Service Provider users in the national capital region, and a literature review.

The report made several observations:

- DoD has not provided a high-quality or reliable end-user experience for all IT users, nor does it routinely measure and share data about the user experience or its impact on worker productivity or mission efficiency.
- There is no single entity responsible for IT services, and this causes inefficiencies.
- Operations and mission-support functions are prioritized over IT support services.
- DoD is struggling to recruit, train, and retain IT talent, which results in significant gaps.
- Supporting enterprise IT requires attention from leadership, reliable levels of funding, and greater coordination across IT departments.

The Defense Business Board found a lack of enterprise IT user-experience performance metrics within DoD. When user experience was measured, the sample sizes were small. Changes were made in a reactive, not proactive manner. The survey found that 80 percent of respondents rated user experience average or below. Respondents expressed concerns about log-on times and frequency of calls to the help desk or submissions of on-line "tickets" reporting IT problems, along with the wait time to have those issues resolved and the frequency of re-authentication. Infrastructure to proactively fix issues is lacking. Service and device management fall under multiple military departments, multiple help desks are not standardized or integrated, and the minimal IT metrics that do exist are not standardized or focused on end-user productivity. End-user device hardware is old, with the average age of desktop equipment being six years. Inconsistent life-cycle replacement plans and an old inventory management system further exacerbate the problem. Budgets do not prioritize IT, and

[1] Defense Business Board, *Recommendations to Improve IT User Experience Within DoD*, February 2, 2023.

bureaucracy hinders life-cycle replacement. Furthermore, configuration images are not standardized due to the wide range (and age) of hardware used across DoD. When hardware is replaced, policies seek the lowest cost and minimally technically acceptable items. Purchasing decisions are made by individual entities and not enterprise-wide.

The report made several recommendations to improve end-user experience, including endpoint monitoring and IT funding prioritization; the use of IT metrics; a review of device replacement and life-cycle management; simplifying security layers and implementing Zero Trust more quickly;[2] the establishment of chief experience officers; centralizing acquisition and negotiations with vendors; streamlining, standardizing, and consolidating help desks; centralizing reference architecture and network and security standards under the DoD CIO; using a federated model; and clearly defining the Defense Information Systems Agency's role for user experience on unclassified systems.

Deferred Investment in Information Technology Infrastructure

In early 2023, the DoD CIO established a task force to assess the impact of investments made to address technical issues resulting from prior deferral of investment in IT infrastructure.[3] The task force was charged with developing outcome-based metrics for IT performance and security to be used as a measure of assessing the impact of IT infrastructure investments; establishing end-user and network performance metrics to support follow-on investment reviews specific to IT infrastructure and user experience; and monitoring those metrics with semiannual updates. The scope of the tasking covered upgrading IT infrastructure on the department's enterprise-wide unclassified and secret-level classified networks and end-user performance. Though each military department and the Defense Information Systems Agency collect metrics relevant to the effort, the task force was not able to establish common metrics across all organizations. Instead, they endorsed 29 metrics specific to different organizations as the initial baseline, with only one (the percentage of end-user devices

[2] A zero-trust paradigm assumes that no actor, system, or network is trusted, even those operating within a security perimeter. DoD, *Department of Defense (DoD) Zero Trust Reference Architecture*, Version 2.0, July 2022.

[3] Although DoD refers to the issues that arise from deferred investments in IT infrastructure as *technical debt*, we do not use that term in this report. This is because the term *technical debt* more commonly refers to the deliberate accrual of unperformed work on an IT system due to rapid software development, which is beneficial in the short term but jeopardizes the long-term health of a system. For more information regarding the meaning of the phrase *technical debt* in software development, see Robert Nord and Ipek Ozkaya, "10 Years of Research in Technical Debt and an Agenda for the Future," *SEI Blog*, August 22, 2022.

DoD's lack of investment in IT infrastructure may be deliberate, but it more commonly arises from a desire to fund emerging technologies and programs in lieu of the sustainment of older systems. Recommended reading regarding the perpetual underfunding of DoD infrastructure sustainment and maintenance (of all types) includes: GAO, *Defense Infrastructure: DoD Should Better Manage Risks Posed by Deferred Facility Maintenance*, GAO-22-104481, January 31, 2022; Patrick Mills, Muharrem Mane, Kenneth Kuhn, Anu Narayanan, James D. Powers, Peter Buryk, Jeremy M. Eckhause, John G. Drew, Kristin F. Lynch, ed. James Torr, *Articulating the Effects of Infrastructure Resourcing on Air Force Missions: Competing Approaches to Inform the Planning, Programming, Budgeting, and Execution System*, RAND Corporation, RR-1578-AF, 2017.

In 2019, in an effort to upend a culture that favors funding new development versus sustaining older systems, the Defense Innovation Board recommended a single color of money to fund software and IT-related programs. Congress has approved a limited number of programs to use this approach but has not yet adopted it. Defense Innovation Board, *Software Is Never Done: Refactoring the Acquisition Code for Competitive Advantage*, May 3, 2019; Jared Serbu and Scott Maucione, "Congress Taps Brakes on DoD Project to Reform IT Funding," *Federal News Network*, March 14, 2022.

compliant with Microsoft Windows 11) common to all.[4] Thus, the military departments, the Defense Information Systems Agency, and the Joint Service Provider (which supports the national capital region) continue to measure and report between six and eight similar metrics for IT infrastructure and end-user performance to provide a picture of progress against those issues for DoD.

The task force made general findings and observations, several of which are also in the Defense Business Board report.[5] These findings and observations include the following:

- There is no consistent enterprise-wide approach to performance monitoring.
- Life-cycle replacement for end-user devices and operating system upgrade practices are inadequate.
- Currently programmed funding for addressing technical debt and user experience is insufficient.
- There is a need for DoD-wide user experience and technical debt management and governance.

These findings resulted in the DoD CIO's recommending that DoD increase funding for enterprise-wide performance monitoring, life-cycle replacement of devices, and targeted IT infrastructure focus areas to remediate many of the problems identified. CIO also recommended establishing a "User Experience and Technical Debt Portfolio Management" office. As of the time of this writing, although that office has been established and staffed, the DoD CIO is still working to secure funding for the other efforts.

Initiatives to Improve Information Technology Performance in the U.S. Department of Defense

Several DoD organizations are already collecting and using user-experience data in various forms to identify technical issues and drive design improvements.

The Department of the Air Force (DAF) established a chief experience officer position in 2019 and soon began deploying digital experience monitoring agents on end-user devices. By the end of 2023, they had agents deployed on about 5 percent of the U.S. Air Force and U.S. Space Force unclassified computers.[6] Data are sent to a central collection system for analysis. At the time of our study, DAF was also regularly tracking and reporting the performance of their unclassified wide area

[4] Department of Defense, Office of the Chief Information Officer, *Department of Defense Tech Debt Metrics Report*, June 2023.

[5] Also in 2023, RAND conducted a study for the Army to assess its aging IT infrastructure and associated risk. The findings of that study are consistent with those of the Defense Business Board (DBB) and the task force. The report also highlights, in a very practical way, how difficult it is to identify, track, and sustain IT equipment, which is a necessary first step in being able to understand the scope of deferred investment in IT infrastructure. Bradley Wilson, Padmaja Vedula, Aimee Bower, Timothy Parker, Giovanni Malloy, Lisa Colabella, Erin Leidy, Ada Ibeanu, Madison Williams, *(U) Aging Systems in the Information Age: An Assessment of Technical Debt in Army Enterprise Information Technology*, RR-A2433-1, August 2024; not available to the general public.

[6] Colt Whittall, "Update: How We Are Fixing Our Computers," presentation, August 28, 2023b, slide 13. Whittall was the Air Force chief experience officer.

networks and base area networks[7] through a performance monitoring system operating on more than 75 of the Air or Space Force bases, with plans to expand coverage.[8] DAF also uses a short survey instrument, called the Air Force IT Pulse, which asks basic questions about user satisfaction in their IT as well as optional questions on the usability and suitability of specific systems from among about 40 applications that have more than 50,000 user accounts.[9] DAF sends the survey to one-twelfth of the total population of DAF users each month (about 30,000 Air and Space Force members), so that every user has a chance to respond to one survey each year. DAF reports an approximately 6-percent response rate to the surveys.[10] In addition to these centrally driven efforts, the DAF chief experience officer has made web analytics available for the use of application development teams. By adding two lines of instrumentation to a website or application's code, the teams can record the website or application's (a) page load times, (b) client, server, or network accesses, and (c) navigation and usage. These metrics are placed in a common repository for performance analysis.[11] The former DAF chief experience officer commented:

> UX [user experience] and performance varies widely by location, audience and application. . . . The AF IT Pulse survey, Aternity, NetScout and other tools in DAF confirm variances of 3× to 5× in key UX metrics that reflect concentrations of old PCs at particular locations, network issues at particular installations, specific applications that are very poorly designed and implemented, etc. The mission impact of IT thus varies substantially and the worst effects can be concentrated precisely where we do not want them. The inverse is also true. UX and performance at some locations, including headquarters, for most users, can be comparable to IT in large commercial organizations. This also represents an opportunity. User satisfaction with the help desk can range 3–5× from the worst to the best installation. There is an opportunity to apply whatever works at the highest rated installations to the lowest rated installations.[12]

DAF is also working to address user-experience principles throughout the system life cycle, offering sample language for requests for proposals and other contract-related instruments and

---

[7] They have also begun to run these tests on some of their classified networks. Conversations with DAF personnel regarding Whittall, 2023b.

[8] For reference, the DAF has approximately 100 bases worldwide, but also supports personnel assigned to U.S. joint and allied bases and installations.

[9] To broadly improve user experience, DAF selected applications to monitor based on the number of users rather than to a more mission-focused metric (Colt Whittall, "How We Are Fixing Our Computers," *Medium*, March 16, 2023a). However, our research does not find that frequency of use is a reasonable proxy for mission criticality. Furthermore, we do not find that applications are the best unit of measure in finding the 10 percent of application usage that leads to unacceptable losses of productivity.

[10] With a response rate of only 6 percent, the representativeness of the sample is in question. As with our survey—which has a very similar response rate—it is likely that people with bad experiences are more motivated to reply to the survey, thereby introducing self-selection bias to the results. From our survey, we know that some users had to persevere through multiple IT-related challenges to access, complete, and submit their responses.

[11] Whittall, 2023b, slide 38.

[12] Colt Whittall, "10/24/24 Colt Whittall Comments re Rand Report 'Underperforming Software and Information Technology in DOD,'" memo to Sarah Zabel, November 6, 2024.

supporting the formation of user-experience communities for designers, developers, and software engineers. A guidance memorandum and governance charter, currently in staffing, would require in-scope software development efforts to adopt web analytics and user feedback throughout the life cycle.

A Department of the Navy (DoN) pilot program moves many functions to a cloud environment, to which users can connect by using a remote desktop application. The "Last Mile Challenge," which is led by the Program Executive Office Digital, seeks to provide "secure reliable commercially available cloud services for ships and bases."[13] DoN is now configuring their end-user devices with fewer applications, bypassing the Navy-Marine Corps Internet when it is not required and instead using a more direct path to the public internet. In many cases, connections from on base are slower than connections off base due to outdated Navy base communication infrastructure. Correcting this may be difficult, as noted in a recent article for *Federal News Network*:

> I think we've done a pretty good job on the shore side of upgrading the off-base transport—that infrastructure is really showing a great improvement," said Skip Hiser, the CIO for the Navy's Fleet Forces Command. "But then we hit the installation, and we have decades-old stuff in the ground. It might be twisted-pair copper, it might be multi-mode fiber, it might be single-mode fiber, it's a variety of stuff. If we try to upgrade all the installations and their infrastructure, it's going to be billions of dollars, which is just not in the budget. So what's the alternative? It might be new construction, it might be hybrid solutions involving fiber and wireless. But it really takes a survey of each individual base to understand that infrastructure and how you're going to upgrade it.[14]

The Navy is also working to fully implement a measurement framework for their IT and software performance, Worldclass Alignment Metrics, initiated by Program Executive Office Digital.[15] Worldclass Alignment Metrics roll up data from over 300 operational metrics to a set of 15 technology outcome-driven metrics, which are further rolled up to five mission outcome-driven metrics. Operational metrics include endpoint performance measurements, service-level requirements, network performance data, help-desk metrics, user feedback from survey instruments, data from financial systems of record, system and security logs, workforce data, and contract data and deliverables. Technology outcome-driven metrics are categorized into network health, service health, endpoint health, labor force, communications, business financial management, transformation, modernization, and system updates. The five top-level mission outcome-driven metrics are user time lost, user satisfaction, operational/cyber resilience, adaptability/mobility, and cost per user. Though a great many of the operational metrics are collected through automated means, some data are collected manually.

The purpose of the Navy Worldclass Alignment Metrics is to drive significant improvement across customer experience, acquisition performance, and alignment to mission needs. A dashboard

---

[13] Jane Rathbun, "IT Pilots Show Promise Toward 'Fix My Computer,'" Department of the Navy, Chief Information Officer webpage, August 1, 2023.

[14] Jared Serbu, "Navy Ready to Start Implementing Fixes to Notoriously Slow Computers," *Federal News Network*, May 19, 2023. We note that the issue of outdated on-base communications infrastructure is not unique to the Navy.

[15] All statements regarding the Worldclass Alignment Metrics come from documentation provided to us by the Navy's CIO office. They include PEO Digital, "PEO World-Class Alignment Metrics 101," undated presentation; PEO Digital, "World-Class Alignment Metrics," presentation, 2023; DoN, "Strategic Intent to Implement World-Class Alignment Metrics," draft memo for DoN CIO signature, pre-decisional, undated.

provides views tailored to an executive level as well as to action officers in various roles. A guidance memorandum, in staffing as of early 2024, would establish the Worldclass Alignment Metrics as a centralized investment decision model with the goal of developing spend plans and prioritizing investment decisions for software and IT based on tangible metrics. For this reason, we evaluated the Worldclass Alignment Metrics as a possible framework candidate as described in Chapter 5.

Though the Army does not broadly and methodically collect user-experience metrics at this time, Army units do collect and report data for 97 metrics of performance for communications, computer systems, applications, and related services. These metrics include items that have been shown to be relevant to user experience in software and IT, such as time to resolve user issues, network availability and capacity, and availability of backup and storage. In 2023, the Army asked RAND to assess risk associated with aging IT equipment.[16] The findings of that study mirror those of the others mentioned above.

OSD's Administration and Management Office and the Defense Information Systems Agency's Joint Service Provider have also been active in seeking to understand and improve the experience of users of IT infrastructure in the Pentagon and national capital region. In 2022, OSD published the results of a study on status and issues with IT support provided to the approximately 22,000 OSD users in the Washington, D.C., metropolitan area.[17] These users account for approximately 60 percent of Joint Service Provider full-support customers. Joint Service Provider also provides transport-only support to additional entities in the Washington, D.C., area. The study involved user satisfaction surveys distributed to IT managers and OSD users as well as listening sessions with service providers, customers, and other organizations involved in management, governance, and funding for OSD IT and an extensive document review. The study resulted in the following findings:[18]

- Lack of clear authorities and governance created a gap in requirements management.
- Repeated IT consolidations for efficiency purposes resulted in significant resource shortfalls.
- Lack of OSD IT service standardization affected performance.

After the release of the study, a CIO for OSD was appointed, and OSD has released an IT Enterprise Implementation Plan.[19] Results of the OSD study and the Joint Service Provider customer surveys were among the key data provided to the Defense Business Board for their subsequent study.

Cost of Department of Defense Information Technology and Software

According to the Government Accountability Office's (GAO's) IT Systems Annual Assessment for FY23, DoD planned to spend about $45.2 billion from FY21 through FY23 for its unclassified IT,

[16] Wilson et al., 2024; not available to the general public.

[17] OSD, Director, Administration and Management, *Achieving "Mission Ready": How OSD's IT Enterprise Can Benefit from Refreshed Strategy, Leadership, and Resourcing*, July 2022.

[18] OSD, 2022, p. ii.

[19] OSD, Director, Administration and Management, *OSD IT Enterprise Implementation Plan: The Initial Steps of the Journey Toward Improved Digital Experience*, February 2023.

which includes 25 major IT systems intended to help sustain key business operations such as contracting, logistics, human resources, and financial management, 723 standard IT infrastructure investments to provide the unclassified computing and network environment these systems depend on, and numerous other programs within the unclassified IT portfolio.[20] The report did not include costs of classified IT systems or infrastructure as that information is not publicly released. Consequently, the costs in the GAO report are a lower bound.

IT and software-based systems are key enablers of effective military operations, but only if they are available, accessible, and usable under operational conditions. The GAO assessment noted gaps in performance monitoring, user training plans, and cybersecurity strategy among the 25 major business systems they studied; these gaps implied potential issues with system usability and adoption. They also noted schedule delays and cost growth, both of which occurred in 12 of the 25 programs. Considered along with the studies and anecdotal evidence above, the report calls into question DoD's ability to deliver effective IT and software-based systems in a cost-efficient manner.

Perhaps more critically, the IT systems that GAO identified for inclusion in its report undoubtedly account for much less than DoD's actual total spending on IT infrastructure and application software. DoD's system of record for IT programs, the Defense Information Technology Portfolio Repository, identifies approximately 4,500 systems and enclaves, thus confirming that there are many more systems than tracked in the GAO assessment. As we will discuss in the next chapter, military department CIOs are frustrated by their inability to identify all IT infrastructure and application software within their departments.

[20] GAO, *IT Systems Annual Assessment, DoD Needs to Improve Performance Reporting and Development Planning*, GAO-23-106117, June 2023.

Chapter 3

Policy and Technical Challenges to Improvement

This study seeks to identify challenges, effects, and potential solutions related to the use of software and IT in DoD. Previous studies, news reports, and social media posts have examined this topic and provided information about the current state of IT, but less has been said from the U.S. military department CIO perspective or about policy challenges. The U.S. military department CIO perspective is important, because CIOs are the ones who must implement DoD policies, review budget requests, and see the "big picture" of IT for their departments.[1] We addressed this gap by conducting semistructured interviews with those CIOs to explore their perceptions of policy and technical challenges hindering software and IT reforms within U.S. military departments. In addition, we aimed to gain insight into the processes used and barriers faced by CIOs as they seek to improve IT infrastructure performance and user experience within their military departments.

In the fall of 2023, we conducted five interviews with current U.S. military service CIOs, one interview with a chief technology and innovation officer, and one interview with a former CIO—totaling seven interviews that represent the U.S. Army, Air Force, Space Force, and Navy, as well as the Marine Corps and DoN. Interviews lasted approximately one hour and were not for *direct* attribution, meaning we do not report names or associate comments with specific individuals. Our interview guide (provided in Appendix A) focused on eliciting CIO perceptions of the following:

- the most pressing IT and software issues the organization is experiencing and how those issues vary across functional areas and within versus outside the continental United States (CONUS versus OCONUS) locations
- how the application software and IT infrastructure needs of their organization are determined and prioritized

[1] Per Title 10 of the United States Code, 2022 edition,

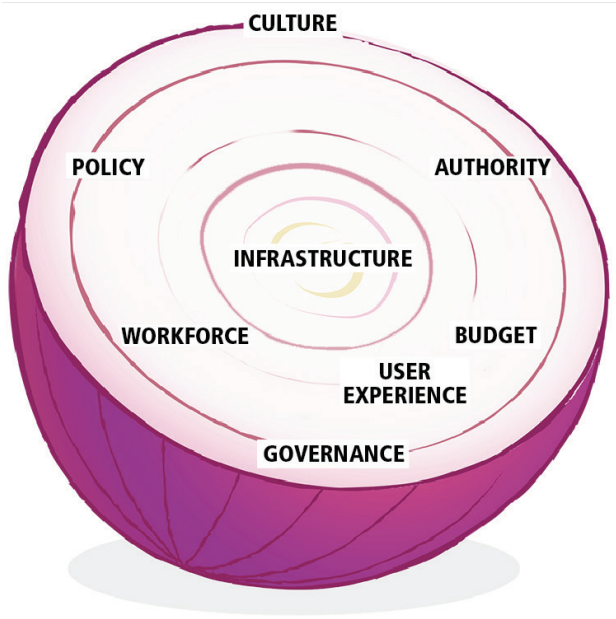
> … the Chief Information Officer of a military department, with respect to the military department concerned, shall—
> (1) review budget requests for all information technology and national security systems;
> (2) ensure that information technology and national security systems are in compliance with standards of the Government and the Department of Defense;
> (3) ensure that information technology and national security systems are interoperable with other relevant information technology and national security systems of the Government and the Department of Defense; and
> (4) coordinate with the Joint Staff with respect to information technology and national security systems (U.S.C., Title 10, §2223 (b). Information Technology: Additional Responsibilities of Chief Information Officer of Military Department).

- how the current state of application software and IT infrastructure could be improved
- policy challenges that inhibit meaningful improvements.

Appendix A provides additional details on the interview and qualitative analysis methodology.

CIOs identified eight challenge areas related to improving the performance of IT infrastructure and application software within their organizations: culture, policy, governance, authority, workforce, user experience, budget, and infrastructure. We conceptually represent the relationships we found between these challenges as the layers of an onion in Figure 3.1.[2] Organizational culture, the outermost layer, had touchpoints in all challenge areas, creating barriers to overcoming policy, governance, and authority challenges, which affected workforce, user experience, budget, and the IT infrastructure at the core of the onion.[3] The next layer—comprised of policy, governance, and authority—was said to influence more central software challenges, including workforce, user experience, and budget; we have located these more central items in the figure in proximity to the challenges they are rooted in. For example, workforce challenges are reported to be rooted largely in policy and governance issues, such as outdated hiring processes and noncompetitive compensation. CIOs discussed user-experience challenges in relation to governance and authority, including slow deployment processes and lack of collaboration. Budget challenges emerged in terms of both how

Figure 3.1. Conceptual "Onion" Representing Policy and Technical Challenges in the Services



[2] The challenges, including how they are organized and relate to one another, are based only on our analysis of the CIO's perceptions as elicited in our interviews. Accordingly, they do not represent our independent view of the policy and technical challenges. We intend the onion analogy to illustrate the interrelated nature of the issues rather than suggesting that outer layers mask inner ones or implying a hierarchy of problem potency. The analogy aims to convey how various issues are interconnected and overlapping, similar to dependencies, rather than being sequentially revealed.

[3] CIOs did not explicitly define *culture*, but we inferred the definition to be the behaviors, beliefs, and traditions of the workplace.

money is spent on IT and software (policy issues) and the process for making those spending decisions (governance issues). Infrastructure challenges, the center, were a consequence of all other challenge areas—that is, how IT and software deployment, management, and investment have been governed, including the disaggregated nature of the IT and software infrastructure.

Culture Drives Some Information Technology and Software Challenges and Can Make Others Difficult to Overcome

CIOs identified organizational culture as a barrier to overcoming policy, governance, and authority challenges and as having downstream effects on workforce, user experience, budget, and infrastructure challenges. Cultural barriers included a tendency for services to maintain status quo technologies, approaches, and mindsets. CIOs stated that current policies allow for decisionmaking regarding IT and software that does not centralize or consolidate efforts. For example, despite the statutory responsibility CIOs have for reviewing budget requests for all IT systems, they do not have acquisition authority. Instead, program executive officers have that authority, and there is no single line item in the budgeting process that captures all of the IT infrastructure or application software requests. CIOs state that program executive officers and individual program offices continue to purchase unique IT infrastructure, even though a shared ecosystem driven by an enterprise services approach may be more beneficial to the services. CIOs also shared that when new policies toward modernization are introduced, a lack of education and experience can lead to slow adoption. CIOs reflected on the challenge of shifting organizational culture, which in turn, makes it difficult to overcome IT and software issues that touch all other challenge areas.

Policy, Governance, and Authority Challenges Are Connected and Pervasive

Policy, governance, and authority are interconnected challenges as discussed by CIOs.[4] While each of these has stand-alone challenges, they can also shape each other. For instance, CIOs described a relationship between policy and authority challenges, such that policy directs how authority can be exercised within the services. Since CIOs' authority (or lack thereof) drives how they can approach governance, the challenges surrounding authority then influence governance. Another visible connection, though not emphasized as strongly in the interviews, is between policy and governance. Policy can affect how governance meets objectives, and governance also drives the creation of policy.

CIOs shared that policy, as written, offers room for many different interpretations and can therefore be applied subjectively. For instance, risk-averse staff may find a policy restrictive, leading to

[4] By *policy*, we mean the processes, structures, and systems that are used to manage and make decisions about IT infrastructure and application software. *Governance* refers to the ability to make informed decisions, and *authority* refers to a set of rules, guidelines, DoD instructions, and so on that delineate the circumstance and limitations under which decisions can be made.

inaction that inhibits progress. In other cases, staff may view policy as a bureaucratic hurdle that can be overcome with reasonable effort by requesting exceptions to policy.[5]

Governance is one way to ensure that the appropriate individuals are involved in creating policies, though CIOs also identified many challenges in governance itself. Governance challenges relate to the structures and processes that inhibit effective decisionmaking—including a lack of awareness across programs and projects—which make it difficult for CIOs to prevent redundancy in effort and investment. Furthermore, CIOs stated that the governance structure does not enable coordination of efforts from headquarters to the edge, nor does it support leadership's ability to federate responsibility at the edge.[6] While CIOs are responsible for the governance structures they set up, perceived limitations to their authority drive how they approach that task.

CIOs identified authority challenges that limit their ability to exercise oversight of the existing IT infrastructure and application software within their organizations. Too often, they find out about issues and decisions after the fact. This means that CIOs are often reactive to IT and software decisions as opposed to being proactive participants in those decisions. These challenges were especially consequential for budget-related decisions. Further compounding this challenge is the external authority required for some decisions, which can result in prolonged processes and actions. For example, the authority to operate process is designed to ensure that changes to IT and software have implemented the National Institute of Standards and Technology's risk management framework and are secure; this is a necessary but often lengthy process that can slow the deployment of improved IT and software.[7] CIOs view the current authority to operate processes as not agile enough for operating in a rapidly changing threat environment and advocate for implementation of tools and processes that would allow for continuous authority to operate. However—despite a February 2022 memo from the DoD's chief information security officer encouraging practices that enable implementers to seek a continuous authority to operate and more recent affirmation of the use of such an authority in DoD's Software Modernization Implementation Plan—the military department CIOs are seeing limited movement toward implementation of continuous authority to operate processes.[8]

Challenges related to policy, governance, and authority were all found to influence other IT and software challenges shared during the CIO interviews. These challenges related to workforce, budget, and user experience, which we discuss next.

---

[5] Exceptions to policy can be obtained by issuing a memorandum signed by a responsible authority stating the reasons for the exception. For use of exceptions to policy by the Navy's Program Executive Office (PEO) Digital in pursuit of more rapid modernization of IT systems, see Darren Turner, "Flank Speed: Exceptions to Policy," *PEO Digital News*, May 11, 2023.

[6] By the *edge*, we mean the lowest echelon that directly uses, manages, and administers the IT infrastructure and application software. *Federation* refers to organizations that use centrally defined standards and shared resources; individual teams have autonomy in implementing those standards.

[7] An *authority to operate* is a formal approval for software or a system to operate on a network. For more on the risk management framework, see National Institute of Standards and Technology, *Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy*, NIST SP 800-37 Rev 2, December 2018.

[8] David W. McKeown, DoD Senior Information Security Officer, *Continuous Authority to Operate*, OSD Memorandum, February 4, 2022. Implementation of a continuous authority to operate is Objective G3.1 Tier 1 of DoD, *Software Modernization Implementation Plan Summary*, p. 6, March 29, 2023.

Workforce Challenges Relate to Policy and Governance Issues

CIOs are concerned about their workforce, especially with respect to hiring, retention, skills, and training. These workforce challenges are rooted largely in policy and governance issues. CIOs described the U.S. Office of Personnel Management's hiring model as outdated and based on policies that no longer serve their intended purpose.[9] They also shared that hiring is slow and incremental. Further, CIOs perceive that lack of flexible hiring options hinders the services' abilities to hire for short-term assignments or to support transitions between government and private-sector positions.[10] CIOs also mentioned that noncompetitive compensation for government positions makes competing for talent with private industry difficult. CIOs viewed this challenge as amplified in regions where government and private industry opportunities are co-located, such as in the Washington, D.C., metropolitan area.

In addition to discussing hiring challenges, CIOs reflected on difficulties retaining talent in the services due to employees' dissatisfaction working with outdated infrastructure coupled with incentives to join industry. As one CIO noted, the reality of conducting duties within an outdated infrastructure becomes apparent to these employees only after hiring. Further, the experience and training gained in the services are attractive to external employers, who can offer higher salaries, more benefits (e.g., on-site childcare), and a more tech-forward working environment. Although these incentives present a challenge for workforce retention, CIOs reflected on the instances in which employees remain with the services for many years or rejoin after leaving; those returning staff attribute their return to personal satisfaction gained from being a part of the services' core mission, which cannot be met elsewhere.

As the services become increasingly digital organizations, CIOs identified a need to evolve workforce skills to match modern-day objectives. Broadening expertise is necessary so that the workforce will have a better understanding across an increasingly integrated digital ecosystem.[11] CIOs suggested that robust reskilling of existing staff will be important, and training will be necessary to develop these new skills. Still, CIOs noted that training has its own challenges, such as training materials not always being aligned to how employees learn best. Additionally, CIOs have lacked capabilities to track the assignment of trained individuals through a talent management system, meaning that matching employees to positions or knowing where skills are distributed can be difficult.

---

[9] For more information about the hiring process model, see Office of Personnel Management, "Hiring Process Analysis Tool," Human Capital Management—Hiring Reform, webpage, undated. The inadequacies of the U.S. government hiring process regarding IT and software talent have been flagged by multiple studies, including Defense Innovation Board, *Department of Defense Workforce: Competing for Digital Talent*, September 15, 2020. Bonnie L. Triezenberg, Jason M. Ward, Jonathan Cham, Devon Hill, Sean Robson, and Jeff Fourman, *The Composition and Employment of Software Personnel in the U.S. Department of Defense: An Initial Analysis*, RAND Corporation, RR-A520-1, 2020. GAO, *DoD Software Acquisition: Status of and Challenges Related to Reform Efforts*, GAO-21-105398, September 30, 2021. We note, however, that there is some controversy regarding the underlying cause of these issues, with some blaming DoD's use of special hiring processes. Jessie Bur, "Have the DoD's Special Hiring Practices Hurt More Than Helped?" *Federal Times*, May 6, 2021.

[10] Note that this perception was found despite the Office of Personnel Management's approval of direct-hire authority for IT management personnel (in place since 2003) and for cybersecurity professionals (in place since 2018). For additional information regarding the Direct Hire Authority, see Office of Personnel Management, "Government-Wide Authority," Direct Hire Authority, website, undated.

[11] Developing and expanding the DoD's digital workforce is Objective G3.3 Tier 1 of DoD, 2023.

Finally, CIOs shared that while they are attempting to hire individuals with expertise that spans multiple IT and software areas, the Office of Personnel Management hiring process lacks specialized position descriptions that address both career field definitions and requirements, which would facilitate these efforts.[12] CIOs reported that this deficiency has resulted in applicant pools that do not fully satisfy the necessary skills that they hope to grow.

Budget Challenges Relate to Authority and Governance Issues

In the CIO interviews, budget challenges emerged in terms of both how money is spent on IT and software (policy) and the process for making those spending decisions (governance). Military department CIOs have statutory oversight of IT and software budgets in their organization, but a lack of centralization and transparency in both budgeting and spending can reduce CIOs' authority, awareness, and ability to govern. For example, since CIOs do not have acquisition authority, IT and software spending can occur without being routed through CIOs' IT governance process; as a result, CIOs expressed that their role becomes more reactive than proactive. This issue is especially apparent to CIOs toward the end of the fiscal year, when functional areas are eager to spend any remaining funds.[13] One CIO shared that this independence leads to spending in IT infrastructure that does not match their guidance.

CIOs also shared that a lack of integrated information makes it difficult to understand the impact of their budget decisions related to IT and software. For example, procurement of application software with its underlying IT infrastructure is often funded as part of a larger system and has not always been adequately segregated within that larger system's accounting. This means that historical procurement data on IT infrastructure and application software are not centralized or easily searchable. Not having the ability to view, filter, and analyze budget-related information makes aligning funding to certain functional areas or developments a challenge for CIOs, especially given the scale of the military departments.

The timeline of DoD's planning, programming, budgeting, and execution process inhibits CIOs' ability to effect change.[14] In the fall of 2023, CIOs reported that they were actively working to establish FY26 budgets for IT and software capabilities, with the budgets for FY24 and FY25 having

---

[12] While there is a career field for IT professionals, there is no similar career field for software professionals. For recommendations to overcome the hiring and training challenges that arise from the lack of a career field and the lack of a competency model to use in describing the skills needed, see Sean Robson, Bonnie L. Triezenberg, Samantha E. DiNicola, Lindsey Polley, John S. Davis II, and Maria C. Lytell, *Software Acquisition Workforce Initiative for the Department of Defense: Initial Competency Development and Preparation for Validation*, RAND Corporation, RR-3145-OSD, 2020.

[13] As noted in Defense Innovation Board, 2019, the line between research, development, test, and evaluation and operation and maintenance funding is often impossible to distinguish for IT infrastructure and application software. Since IT-related hardware and software items or related services are relatively low cost and widely available for purchase, they represent a relatively easy way for DoD organizations to obligate remaining funds of either color of money at the end of a fiscal year.

[14] Due to DoD, executive branch, and congressional appropriations constraints, the military departments make their inputs to the planning, programming, budgeting, and execution process two years in advance of the year in which an appropriation is made. While there are ways to make budget requests outside the cycle of this process, they are rarely granted. For more information, see Congressional Research Service, *DoD Planning, Programming, Budgeting, and Execution (PPBE): Overview and Select Issues for Congress*, R 47178, July 11, 2022.

been established in prior years. While all military acquisitions are subject to the planning, programming, budgeting, and execution process, the long lead times it creates may be particularly detrimental in areas (such as IT) where DoD would benefit from leveraging rapidly emerging technologies.[15] CIOs noted that it is extremely challenging to anticipate IT vendor pricing years in advance or to anticipate rapidly evolving IT and software needs.

One final budget challenge reported was the lack of priority for IT and software within DoD's budgeting process. CIOs shared that funding has historically been prioritized to weapon systems, incentivizing a practice of funding upgrades to IT infrastructure and application software as part of those weapon systems. This practice results in insufficient funding for basic computing infrastructure, such as wide area networks and personal computers that are not thought of as being part of a weapon system or other program of record (and, as noted above, this practice fragments funding for IT and software, which makes it harder for CIOs to provide oversight). Further, the funds received for IT and software through this process are often not spent in an integrated way. One CIO noted that such an approach leads to the services losing negotiation power as multiple Program Executive Offices and program offices individually purchase the same service from a vendor.

User-Experience Challenges Relate to Governance and Authority Issues

Unlike the other challenges identified, CIOs generally did not initiate discussion of users' experiences, but instead, shared thoughts on this topic when prompted. Multiple CIOs expressed that they did not believe that current IT infrastructure and application software were always meeting user needs. One area of dissatisfaction centered on legacy IT architecture, especially with regard to network operations. One CIO shared that efforts to centralize and optimize network services would improve user experience. Other CIOs shared that implementing IT infrastructure upgrades and software capabilities is a slow and lengthy process. They noted that unlike in the commercial world, the services have a strong risk-averse culture that requires extensive vetting of new capabilities, which involves balancing the need for security with unique operational challenges. CIOs believe that while this thorough process is important for maintaining security, it is often too slow and acts as a barrier to deploying capabilities at pace and at scale.

Another area of concern was the extent to which capabilities match user needs. End users are not always involved in the services' acquisition processes. A lack of collaboration between designers, decisionmakers, and end users means that new IT and software capabilities may not address users' wants or needs.[16] CIOs reflected on how system design can be personality driven, such that capabilities

[15] Reform of the planning, programming, budgeting, and execution process is a topic of concern for both DoD and Congress. In the FY22 NDAA, Congress established the Commission on Planning, Programming, Budgeting, and Execution Reform to outline a path forward. The commission released its interim report in August of 2023. Commission on Planning, Programming, Budgeting, and Execution Reform, *Interim Report*, August 2023.

[16] This complaint is not unique to IT infrastructure. Although modern software development processes emphasize early engagement with users, DoD struggles to find a way to prioritize that engagement. GAO first surfaced this problem in 2019 and reported in 2022 that it continues to be an issue. See GAO, *DoD Space Acquisitions: Including Users Early and Often in Software Development Could Benefit Programs*, GAO-19-136, March 18, 2019; GAO, *Leading Practices: Agency Acquisition Policies Could*

are developed based on an individual's ideas rather than executed with a clarity of purpose from headquarters to the tactical edge. This approach can lead to the implementation of capabilities that do not match service requirements or that do not have long-term sustainment plans. CIOs provided multiple instances in which loss of an individual developer resulted in loss of support for delivered capabilities that relied on that individual.

Despite challenges identified surrounding user needs, CIOs recognized the importance of incorporating users into an agile system-design process and discussed ways to improve the user experience by conducting in-person engagements to understand users' needs out to the tactical edge; measuring user experience to better track the performance of IT and software; identifying service personas and their IT and software needs; and budgeting accordingly. Several of our recommendations in Chapter 6 echo these thoughts.

Infrastructure Challenges Relate to All Other Challenges

All challenges identified by CIOs feed into the infrastructure challenges that are at the core of Figure 3.1. Addressing infrastructure challenges requires considering and addressing aspects of all other challenges.

CIOs observed that the at-risk status of IT infrastructure and application software across the services is a consequence of how IT and software deployment, management, and investment have been governed. CIOs reported that their infrastructure depends on old technology, which is failing at an alarming rate and in some cases, without a phase-out plan. Further, this old technology inhibits the adoption of modern capabilities into the infrastructure.

A theme across the CIO interviews is the disaggregated nature of the IT infrastructure and application software, especially with respect to data management and digital integration. For example, systems are designed and developed separately with no forcing function to encourage commonality; logistics and supply chains are treated uniquely; and data are not fused or centralized. Infrastructure is not integrated, although, according to CIOs, stovepipe systems should not be used. They viewed the diversity of the technical stack as too complex, resulting in unforeseen consequences when changes are made.[17] CIOs mentioned the need to flatten networks and centralize services, which they expect will allow IT staff to spend more time on other pressing challenges.

CIOs are interested in moving to the cloud, but they note that some nuance is necessary. For one, they say a command and control mechanism to monitor and determine the behavior of the cloud itself is necessary. Security of the cloud is also an important concern. CIOs assert that enterprise contracts should be used when purchasing cloud services, preventing duplicate purchases and improving management and security.

CIOs discussed the need to continuously monitor the network for threats. The commonly used authority to operate process is not sufficient for cybersecurity, as it grants permission without

---

Better Implement Key Product Development Principles, GAO-22-104513, March 10, 2022. A 2020 RAND report that examined issues in DoD software acquisition also found lack of early engagement with users to be a concern; see Triezenberg et al., 2020.

[17] We note, however, that a centrally managed homogenous system should not be the goal. Diversity in the infrastructure adds resilience to attack or common cause failures, meaning that a balance between commonality and diversity is needed.

continuous monitoring, CIOs argued. Instead, devices that connect to networks should be continuously checked to ensure that they adhere to security policies.[18]

CIOs expressed a pressing need to integrate systems and spending to build operational architecture, since many aspects of the digital ecosystem must operate together in a specific manner to be successful (e.g., commercial technology, cybersecurity, and the cloud). Further, CIOs stressed the importance of dependable IT infrastructure and application software. For example, ensuring data flow, storage, and classification mechanisms are integrated to allow for decision-ready capabilities at the edge is essential to warfighters. CIOs stressed the need to collect and process data in real time and making data (some of which might not be digitized) available to users. CIOs shared that achieving successful integration will require improved governance of IT and software spending decisions—such that decisions are made as a service rather than as a program, and investments can be leveraged across the service (e.g., one contract for one IT service through enterprise strategic sourcing).

[18] While this perspective aligns with the views commonly expressed by CIOs, it is important to acknowledge that it is not without its challenges. Balancing security requirements with usability is a well-documented issue in numerous reports, including sections of this report. Overly stringent security measures can sometimes impede user experience and operational efficiency, necessitating a careful consideration of trade-offs between security and usability.

Chapter 4

Measuring User Impacts and Satisfaction

In this chapter, we provide an overview of the survey instrument and choices we made in formulating the survey questions. We then apply what we learned from the survey to address IT infrastructure and software impacts to DoD's user productivity, mission readiness, and user satisfaction. Appendix B includes the survey instrument and detailed question-by-question review of the survey results.

Survey Design

We constructed our survey to solicit information to answer the first three of our research questions:

1. How much time does it currently take for DoD personnel to access DoD IT infrastructure and applications to begin productive work?
2. Once productive work begins, how much effort is spent resolving issues or recovering effort lost when IT infrastructure or application software malfunctions?
3. How satisfied are DoD personnel with the IT infrastructure and application software supplied to them to perform their mission-essential tasks? To perform other tasks?

To avoid survey fatigue, we limited the number of questions (our goal was that the survey can be completed in 20 minutes or less for the average respondent) and provided multiple-choice answers when possible. In total, the survey consists of 19 questions, nine of which can optionally be repeated for up to six software applications of the respondent's choice. The last question is open-ended and provides a means for survey respondents to share their pressing issues.

To avoid saliency bias, we deliberately did not ask respondents to report "typical" values for log-on, application start-up, and recovery times.[1] Instead, we asked them to report on their most recent attempt to log on or start an application or the most recent time when they needed to resolve an issue and then followed up with a question as to whether the reported experience was better or worse than typical. In all cases, the majority of respondents replied that the reported experience was "about the same," with an equal portion reporting the experience was better or worse. Therefore, we are reasonably confident that the time-based results reported here are unbiased by saliency effects.

[1] By *saliency bias* we mean the "tendency to focus on items or information that are more noteworthy while ignoring those that do not grab our attention" (Decision Lab, "Why Do We Focus on Items or Information That Are More Prominent and Ignore Those That Are Not?," blog post, undated). In this case, we hypothesize that particularly bad experiences are more salient and will bias respondents' reporting of what is average or typical.

In constructing the multiple-choice responses, we limited the number of possible selections to five to avoid fatigue and used time frames that are easy to recognize (e.g., "less than 40 hours a month" since most workers have an intuitive feel for the duration of a 40-hour work week). In general, the first choice represents what we consider "good" responsivity of a system (and therefore no impact on user productivity), and the last choice represents what we consider extremely unacceptable responsivity. Between those two extremes, we provide breakpoints that represent meaningful steps in technical performance improvement. For example, the possible choices for our question "For your most recent logon to the network identified in question 2, how many minutes did it take from when you began the process until meaningful work could be started?" include

- less than 1 minute
- between 1 and 5 minutes
- between 5 and 15 minutes
- between 15 and 30 minutes
- more than 30 minutes and/or I was unsuccessful.

In this way, we prevented later interpretation of trends in the technical performance data that would assess an improvement from 30 minutes to 20 minutes (in our view, a welcome but less than meaningful improvement) equally to an improvement from 10 minutes to less than a minute (in our view, a very meaningful improvement).

For research question 3, we used an industry standard measure for evaluating user satisfaction with software applications called the Usability Metric for User Experience (UMUX)-Lite.[2] This comprises two of the nine questions asked for each reported software application. The two questions are "Please indicate your level of agreement with the following statements: '[Application name] meets my requirements,' and '[Application name] is easy to use.'" Agreement was on a five-point scale from strongly disagree to strongly agree.

Survey Analysis Results

Appendix B provides the survey instrument and an analysis of the responses to each question. In this chapter, we provide synthesized results for the research questions listed above and discuss how our analysis has shaped our understanding of the impact poorly performing IT and software has on productivity, mission readiness, and user satisfaction.

Networks and Network Access

Network Quality. While users generally rate their unclassified networks as high speed, when given a choice between describing their network access as high quality versus intermittent, half characterized it as intermittent. This result holds across all services with civilians significantly more likely (70 percent versus 55 percent for the uniformed military) to characterize their access as high

[2] James R. Lewis, Brian S. Utesch, and Deborah E. Maher, "Investigating the Correspondence between UMUX-Lite and SUS Scores," in A. Marcus, ed., *Design, User Experience, and Usability: Design Discourse*, Springer, 2015.

speed and slightly more likely (55 versus 50 percent) to characterize it as high quality. Notably, the 23 percent of our respondents who indicated they had logged on to the network remotely on the day they answered the survey were significantly more likely to characterize their unclassified network access as high quality (61 versus 49 percent), indicating that *commercial network service providers are perceived as providing better reliability than DoD's on-premise networks*.[3]

Network Log-On Experience. Overall, half of all respondents reported that the time it takes from when they begin the process of logging into DoD systems until meaningful work can be started is under 1 minute, which we characterize as good. Only 13 percent reported that it takes more than 5 minutes. However, a small subset within that group (4 percent) reported that it takes more than 15 minutes to access their systems and begin meaningful work. For many of them, this involves multiple restarts of their computers. For those users, this lengthy access time is a source of great frustration. As one user told us,

> At least three times a week, it takes between 24 minutes and 64 minutes for me to be able to start using my [DoD-supplied] laptop from the time I try to log in. I've started keeping a log because the wait time is so absurd. . . . Eighty percent of the time, my computer starts up with a black screen and no icons. I constantly have to restart my computer because it is non-responsive. . . . [. . .] It is so bad, I don't see how it is not a national security issue.

The results for system access time do not significantly change by service affiliation, network, or whether the respondent is based OCONUS, which makes it difficult to find the small subset of users who suffer these long (and unacceptable) wait times.

Remote access users reported slightly longer access times, but the average additional time is less than one minute. In return for this small additional access time, users experience the higher reliability of the non-DoD networks—a trade that appears to be well worthwhile given that the higher reliability means that they will not have to log on as often. Although it is unclear whether these external networks will be able to sustain that higher reliability performance when under attack, remote access appears to improve productivity for DoD in the current environment.

OCONUS users of classified networks also reported slightly longer access times, with the average additional time approaching two minutes. Ten percent of OCONUS secret-level internet protocol router users reported that it takes longer than 15 minutes to gain access and begin meaningful work. This is unacceptably long and, at 10 percent of the OCONUS force, is likely to affect mission readiness. *Therefore, when considering where to make investments in network infrastructure, we recommend that secret-level internet protocol router circuits OCONUS be prioritized.*

There is one other category of users who experience unacceptably long access times: those users who need to log on for only a few minutes to perform a routine task and do so via a shared laptop or desktop. We identified this class of users from our analysis of their open-ended comments and thus cannot quantify their prevalence in the sample. A significant portion appear to be reservists. These users complained of lengthy durations to log on, which many of them attributed to the large number

[3] Network ratings for remote secret-level internet protocol router users are also slightly better than ratings from those who access it directly. We did not have a sufficient sample size for remote Joint World-Wide Intelligence Communication System users to generate a valid comparison.

of user profiles stored on the laptop or desktop. Others complained more generally that the computing resources are consumed by "bloatware"—applications that are unnecessary for their assigned duties and that slow responsiveness. For these users, even the average time lost when logging on, starting applications, and recovering from issues exceeds their tolerance level. They just want to get on the computer, get off, and continue with their assigned duties. As one said, "This isn't my job. Fix it." *We concur and recommend that DoD conduct periodic reviews of the standard configuration and provide tailored configurations that minimize start-up time for users of shared laptops.*

Virtual Private Network Robustness. Just prior to and during the time of our survey, two of the services had decommissioned their primary virtual private networks (VPNs) after finding security flaws in them.[4] VPNs were the sixth most frequently rated mission-critical application in our survey, and both the absence of a VPN solution (for those who had lost access) and the frequency of VPN dropouts (for those who still had an approved VPN solution), figured prominently in respondents' open-ended comments. VPN substitutes, while they may provide remote access, often restrict user access to sensitive applications and data, making them less useful for mission-critical work. VPN dropouts cause significant downtime and hamper users' ability to complete essential tasks. Therefore, we recommend that DoD improve the robustness and resilience of their remote access offerings.[5] *At minimum, military department CIOs should have at least two VPNs (or other remote access solution) certified for use, with a continuity of operations plan in place to deal with loss of one.*

Application Performance and Issue Resolution

In the survey, we provided the opportunity to rate three applications "most critical to the performance of your duties" and then three applications that "inhibit productivity." We use these distinctions when providing summary results in this section. However, there is considerable overlap between the two groups: four of the five most cited mission-critical applications are also four of the five most cited productivity-inhibiting applications.[6] An application that is mission critical to one user need not be to another. The primary distinction between the two groups is how often the respondents use the application during their duties; half of all mission-critical applications are used for more than 50 percent of the respondents' assigned task, while those applications cited as productivity inhibiting are used less often. Thus, we assume that time lost when using mission-critical applications has a direct impact on mission readiness (i.e., any outage will be difficult to work around) as well as on productivity. Furthermore, we assume that the impact of time lost when using applications respondents classify as productivity inhibiting is primarily on productivity, not mission readiness.[7]

---

[4] Jason Miller, "Navy Used Threat of Cyber Vulnerability to Expand VDI," *Federal News Network*, February 16, 2024.

[5] In addition to the Navy's loss of their primary VPN, the Air Force had also run into issues with their remote access offering at the time the survey was conducted.

[6] We removed duplicate entries from a respondent. Thus, this overlap is not a function of double-counting.

[7] This is a simplifying assumption and does not mean that time losses associated with productivity-inhibiting applications do not have an impact on mission readiness, just that those losses are easier to work around.

Application Start Times. Ideally, applications should start up instantaneously or in less than a minute.[8] However, respondents reported that over one-third of mission-critical applications and nearly one-half of problem applications take more than two minutes to start up. On average, mission-critical applications take from 1.5 to 3.5 minutes to start up, and problem applications take from two to 4.5 minutes to start up.[9] For those who experience multiple issues per day with an application, these start-up times can quickly become a source of frustration.

Application Issues Encountered. Interestingly, there is little difference in the type of issues encountered in the last month when comparing ratings for mission-critical applications with those that are productivity inhibiting. While having an application hang is the most reported issue (~45 percent of applications), lack of accessibility, failures to open, and unexpected closures are nearly as prevalent at ~35 percent each. Issues that result in a loss of data are less prevalent—at around 20 percent—which is not surprising given that Microsoft Office applications are the most frequently rated applications, and these applications routinely provide auto-recovery files when closed unexpectedly.

Almost half of all applications rated by the survey respondents have encountered multiple types of issues in the last month. Note that we did not inquire as to the frequency with which these issues occur, but from the open-ended comments we know that for some users severe issues are experienced multiple times a day. This high percentage of applications with rather severe issues exceeds most norms for software application reliability and is undoubtedly a factor driving perceptions of the inadequacy of DoD IT and application software.[10]

Issue Recovery Times. One-quarter of both types of applications take more than 15 minutes to recover from an issue, a percentage that we consider unacceptable when combined with the high percentage of applications that have multiple issues per month. Within those applications that take longer than 15 minutes to recover, we see significant differences between the mission-critical and the

---

[8] Our review of the literature on human-machine interactions indicates that a computer's response time of a tenth of a second feels instantaneous to the user. One second feels to the user as if they are operating the machine directly with no interruption in flow of thought for the user, but a delay of 10 seconds interrupts users' thought processes. Clearly, the average application start-up times experienced by DoD users can interrupt flow of thought. This disruption may not affect mission or productivity if applications are started only at the beginning of a shift or if the delays are expected and the user adjusts to them. However, even when delays are repeatable and expected, research has found that delays that impede the necessary pace of the task at hand will result in productivity loss and user dissatisfaction. For more information on human-machine interactions times, we recommend the original research by Robert Miller, "Response Time in Man-Computer Conversational Transactions," *Proceeding of the AFIPS Fall Joint Computer Conference*, Vol. 33, 1968. For a more contemporary discussion—including the need to pace computer interactions to the pace of the task—see Dabrowski and Munson, 2011.

[9] The high end of our estimate for average start-up time of the productivity-inhibiting applications is undoubtedly low. The large number of responses in the "more than 5 minutes" category means the distribution of responses is neither a bell curve nor uniform. A strictly algebraic computation of the average (or mean) is therefore invalid, producing an estimate that is too low. Future survey designs should add a "5–10 minute" category, with the final choice being "more than 10 minutes."

[10] The advent of application storefronts has provided a means to collect data regarding users' tolerance for poorly performing applications. Based on this data, several application developer websites state that the benchmark for a "good" or "3-star" application is fewer than 1 percent crashes per user, <0.1 percent crashes per session, and <0.01 percent crashes per screen view. See, for example, Gary B, "App Crash Research: Is Your App below or Above the Benchmark," *Medium*, Aug 17, 2020. Assuming that mission-critical applications are in use 20 sessions per month, the average application in DoD experiences more than 0.1 percent severe issues per session. Future surveys should inquire as to the frequency of these severe issues to better quantify this metric.

productivity-inhibiting applications. A higher percentage of productivity-inhibiting applications takes more than an hour to recover (9 percent versus 7 percent); 10 percent of productivity-inhibiting applications never recover at all (as compared with 4 percent for the mission-critical applications).

Lost Work Regeneration Times. The time needed to regenerate lost work also distinguishes mission-critical applications from productivity-inhibiting applications. While 83 percent of lost mission-critical work can be recovered in less than eight hours, only 75 percent of the work lost when productivity-inhibiting applications have issues can be recovered in less than eight hours. In fact, when it takes more than 40 hours to regenerate work lost, it is twice as likely that the application with technical issues that caused the loss is rated as productivity inhibiting.[11]

Understanding Impacts on Productivity

At the start of our research, we were reluctant to characterize time spent accessing networks and software applications as a direct loss in productivity. If these times are repeatable and known, personnel may be able to fill that time performing tasks that do not require network or computing system access. However, from the comments provided in the survey responses, these times do not appear to be repeatable, and users told us that the reported times include many retries; in other words, users are actively involved in the recovery. Therefore, we treat log-on time, application start-up time, and issue-resolution time as direct impacts on productivity, rather than derating them to reflect multitasking. Time spent regenerating a product that is lost due to IT infrastructure and application software malfunction, which our survey also asks about, is more directly tied to a loss of productivity.

To avoid survey fatigue, we did not ask users how many times they log on per day or how often application issues occur. Without this data, any computation of lost hours per month is only a lower bound. To compute that lower bound, we assume that users need to repeat their network log-on and application start-up when an application has a significant issue. Furthermore, we assume that when an issue is encountered, users spend the time to recover from it and regenerate lost work when necessary. Therefore, for each of the application ratings we received, we compute the minimum hours lost per month as a function of system log-on time, application start time, the number of issue types encountered in the last month, and the time to regenerate lost work. We then compute the mean and distribution of those scores. The average time loss per application is two hours per month.[12]

To convert this to a metric of productivity loss requires that we know the number of applications used per month by DoD personnel. We know that half of our respondents cited three applications as critical to their duties, but we also know that respondents who took the time to answer the survey are more likely to depend on IT and software for their assigned duties and are unlikely to represent the "average" user. Therefore, making the very conservative assumption that an average DoD user needs

[11] This could be interpreted as a good news story in that mission-critical applications are less likely to take more than 40 hours to regenerate lost work. Presumably, users correctly prioritize the regeneration of lost mission-critical work and have created tools and processes to help them do so quickly.

[12] For mission-critical applications, the average time lost is 129 minutes per month. For productivity-inhibiting applications, it is slightly higher, at 164 minutes per month. For a conservative lower bound, we use 120 minutes as the average (two hours). Mission-critical applications have a wider distribution of scores than that seen for productivity-inhibiting applications.

only one application per month, we find the lower bound for productivity losses in DoD due to IT infrastructure and application software issues is

- two hours per month for the average user
- eight or more hours per month for 10 percent of users
- 40 hours or more per month for 1 percent of users.

At the equivalent of a 40-hour workweek, lost productivity due to IT and software issues equates to about 1 percent for the average user, but for 10 percent of users it is a 5-percent loss, and for 1 percent of personnel it is a 23-percent loss of productivity. This 1 percent of users will be very difficult to find since they are spread across the force, are on many different networks, and use a wide variety of applications.[13]

To convert this metric to cost, we need to understand if the applications in use at the time of the outage are more or less likely to be used by a specific rank or salary grade. They are not the most often rated applications (true for both mission-critical and productivity-inhibiting applications), as those in common usage: Microsoft Word, Excel, PowerPoint and Teams, Adobe Acrobat, VPNs, and browsers. The second tier of commonly rated applications are those used less frequently—the training system, the travel system, the major enterprise resource planning systems—but are not uniquely used by any particular rank/salary. Reservists and those who do not use DoD technology often—and thus need to share computers—are particularly hard hit by productivity losses. For them, technology issues too often turn a five-minute task into a two-hour task, and since they all share a computer, they also share the impact.

Therefore, we make the simplifying assumption that the two-hour loss of productivity per month applies to the total workforce and simply take 1 percent of DoD dollars spent on personnel. This rough-order estimate is $2.5B in FY2023.[14] While this metric of cost can be used in return-on-investment calculations, we caution that it should not be used to evaluate trends over time.

Understanding Impacts on Mission Readiness

Impact on Mission Availability

The outage of two hours per month discussed above is for the average application as experienced by the average user. Two hours of outage per month is unacceptable for many weapon systems, yet too many of today's weapon systems do depend at least tangentially on common applications such as

[13] Users who lose more than 40 hours per month experience that loss while using common applications that others report having minimal issues with. Mission-critical applications reported as contributing to 40 hours or more of time lost per month include everything from a troublesome VPN to MHS Genesis.

[14] Our estimate of personnel costs in FY2023 comes from OSD estimates for the FY24 budget submittal. Military personnel costs come from OSD Comptroller, "FY 24 Budget Submittal, Exhibit M-1," undated-a, and do not include contributions to retiree medical health funds. Civilian personnel costs are estimated at 20 percent of the operation and maintenance allocation in OSD, Comptroller, "FY 24 Budget Submittal, Exhibit O-1," undated-b. Usage of 20 percent of the operation and maintenance allocation as a rough-order-magnitude estimate of civilian personnel cost is in alignment with Congressional Research Service, *Department of Defense Budget: An Orientation*, Nov. 2021.

Microsoft Excel or Access.[15] Furthermore, our respondents were quite vocal regarding mission dependencies on Adobe Acrobat for digital signature of orders and other vital communications throughout DoD. Email, chat, and shared drives are used for mission-critical communications. Users reported that their mission-critical applications are in use throughout the duty period, which makes any outage a mission outage. For the 10 percent of users who lose up to eight hours per month of mission-critical work due to IT and software issues, this impact is substantive. DoD must prioritize efforts to identify and improve the experience of the 10 percent of users for whom mission readiness is unacceptably degraded by IT infrastructure and software issues.

As we noted in the background chapter, some of the military services have begun to collect quantitative metrics routinely and automatically from end-user devices to understand and quantify productivity and mission impacts of poorly performing IT and software. One limitation of the service data collection efforts is that they can collect only on the networks and applications they can instrument, which may or may not be the networks and applications that are essential to mission readiness.[16]

For our survey, we asked users to identify the applications most critical to the performance of their duties. When we examine the list of applications respondents cited as mission critical, we find that only 36 applications (or family of applications) received more than 20 ratings, and these make up two-thirds of the total number of ratings received. The last one-third of the ratings are for the hundreds of applications with fewer than 20 ratings. This means that current service efforts focused on measuring the performance of the top 20 or top 40 most-used applications are likely missing one-third of the mission-critical impacts. Furthermore, when we examine the lists of applications where users report the *worst* performance or *worst* user satisfaction, we find that poor performance and user dissatisfaction are rarely a function of the specific application (with notable exceptions discussed in Appendix B) but are instead a function of the environment in which applications run. That is, the difference between users who experience fast start-up and no critical issues with their mission-critical applications and those who experience 8 or 40 hours of outage is due not to the application itself, but to the environment—the combination of hardware capabilities, security tools and policies, communications latency, and other technical factors—the application runs in. These observations lead us to the following recommendation: *When designing automated collections of IT performance data, services should emphasize finding the worst 10 percent of computing environments rather than finding poorly performing applications.* We caution, however, that simply identifying the computing environments does not mean that remedying those environments will be easy. The totality of our research has

[15] Two hours of outage per month is equivalent to an availability of 99.7 percent. Eight hours of outage per month is equivalent to 98.9-percent availability. Forty hours of outage per month is 94.5-percent availability. If the per-user outage times we calculate from the survey mission-critical application data were to be included in weapon system availability calculations (which are computed using outage times that are the square root of the sum of squared independent factors), we suspect poorly performing IT infrastructure and software would be the dominant factor.

[16] By *mission readiness*, we mean the ability of the U.S. military to fight and meet the demands of assigned missions. Our definition reflects that expressed in Congressional Research Service, *The Fundamentals of Military Readiness*, R46559, October 2, 2020. Note that this definition is slightly broader than that of *operational readiness*, which DoD defines as the ability to perform the missions for which a unit, system, or equipment is organized or designed. Joint Chief of Staff, *Terminology Repository of DOD Issuances*, Version 14, September 15, 2023. We have chosen the broader definition because we are asking users about the IT infrastructure and application software needed to perform their assigned duties, which may not always match the duties for which they were organized, trained, and equipped.

convinced us that poor performance is most often caused by a combination of factors, many of which are out of the control of a program office or of a military department CIO.

Impact on Retention and Morale

Mission readiness is affected not just by time lost to IT and software issues. It can also be affected by morale, or lack thereof. To better understand how IT and software issues affect morale, we coded the open-ended comments we received for sentiment and feelings of agency. Annoyance (at 24 percent) and frustration (at 50 percent) are the dominant emotions found among those who commented. Half of those who commented exhibit signs of both strong negative sentiment and a lack of agency (the ability to affect the issues), signaling an emotional exhaustion that could weigh significantly on their decisions to leave or remain in the service.[17] This is a sizable percentage of the respondents who left comments. When translating it to the percentage of the workforce at risk of departure, we need to account for the fact that participants in the survey self-selected into the sample and do not represent average users. In addition, the ~30 percent of respondents who commented self-selected again by electing to tell us about their experience perhaps in a last-ditch effort to have an impact. Therefore, these results do not mean that ~15 percent (i.e., half of the 30 percent of respondents who commented) of DoD personnel are at risk of departure. Accounting for self-selection bias is not a science, but based on the technical barriers we know respondents had to overcome to respond to our survey, we assume only the most motivated did so and discount this measured result by a factor of three when applying to the total workforce. *Therefore, we conservatively estimate that 5 percent of the DoD workforce may be strongly motivated to depart from service due to poorly performing IT and software.*

Appendix B explains our methodology for coding the open-ended questions to better explore the perceptions that give rise to these emotions. The primary coding is for the perceived cause of the emotion (e.g., hardware, software, policy, and so on); a secondary coding concerns the perceived impact to mission or to productivity. For the primary coding, software inadequacies are the most cited perceived cause of technical issues, at 38 percent. These inadequacies include nonintuitive interfaces, frequent application crashes, and the inability to access critical software tools. Hardware limitations are cited almost as frequently as software limitations, at 32 percent. Outdated equipment, insufficient random-access memory, and slow performance are recurrent themes. One commenter summed it up succinctly: "The computers are simply slow. They are slow to boot up, slow to open files and slow to access networks." Respondents told us multiple stories of how slow response times cause a cascade of errors that ultimately prevent them from successfully accessing or using the applications they need. Many respondents told of having to use their personal devices to accomplish basic tasks within a reasonable time frame.[18]

Some users correctly perceive that it is an imbalance between available hardware resources and the demands of the software hosted on that hardware that drives technical issues. When naming productivity-inhibiting applications, 40 responses simply cited 'bloatware' rather than enumerating

[17] James W. Moore, "What Is the Sense of Agency and Why Does It Matter?" *Frontiers in Psychology*, Vol. 7, August 29, 2016.

[18] Ironically, if security applications are a primary cause of slow computing and drive personnel to use personal devices instead, the net impact of adding security applications may actually be making DoD less secure.

specific applications. In general, these users complained that the software on their machines is not tailorable to their needs and that the result is unneeded applications running in the background and slowing all other applications.[19] Of particular note, we found that for users whose duties require use of a shared computer for only a few minutes per day, the average time lost when logging on, starting applications, and recovering from issues exceeds their tolerance level. This illustrates that measured "time lost" may matter less to morale than perceptions of whether those times are acceptable for the mission. In fact, research indicates a mismatch in the pace of computer-assisted operations and the pace needed for the task at hand is a leading cause of user annoyance and frustration.[20] Our secondary coding results are supportive of this finding. Users who expressed annoyance and frustration were much more likely (60 percent versus 20 percent) to cite the impact on their time and the impediments to mission (lack of access, tools, and interoperability) than they were to cite difficulty in use or lack of suitability to the requirements (see Figure B.12). Furthermore, their complaints are not trivial. For example, when commenting on lack of access to websites, lawyers noted that they were blocked from case law databases, medical professionals that they were blocked from researching patient symptoms and treatments in medical journals, Army officers that they were blocked from Army websites, and airmen that they were blocked from Air Force websites.

Understanding Impacts on User Satisfaction

As noted earlier in this chapter, our survey uses the UMUX-Lite scale to measure user satisfaction. The first statement we ask respondents to rate, "[Application name] meets my requirements," relates most directly to mission readiness—that is, whether the application is fit for purpose. The second statement, "[Application name] is easy to use," relates most directly to efficiency and productivity, or whether the application is fit for use. Together, they can be summed and scaled to produce a composite score between 0 and 100 that is directly comparable with software industry performance metrics for "user satisfaction." Interpreting that comparison, however, can be nuanced.

Despite being a scale from 0 to 100, the UMUX-Lite composite score should not be interpreted as a percentage. In 2009, researchers took over 3,500 user satisfaction scores and correlated them to people's subjective rating from "worst imaginable" to "best imaginable" to produce a grading score. They found that systems with scores less than 50 are generally viewed as "not acceptable," between 50 and 68 as "marginal," and above 68 as "acceptable."[21] While we use these ranges to discuss survey results, we and other researchers caution that metrics of user satisfaction (and usability more generally) should be compared only relatively and within context.[22] Clearly, the context of a weapon system is different from the context of the military pay system, and we might expect that a score that

[19] This lack of user tailor-ability also contributes to feelings of a lack of agency, which contribute to emotional exhaustion.

[20] Dabrowski and Munson, 2011.

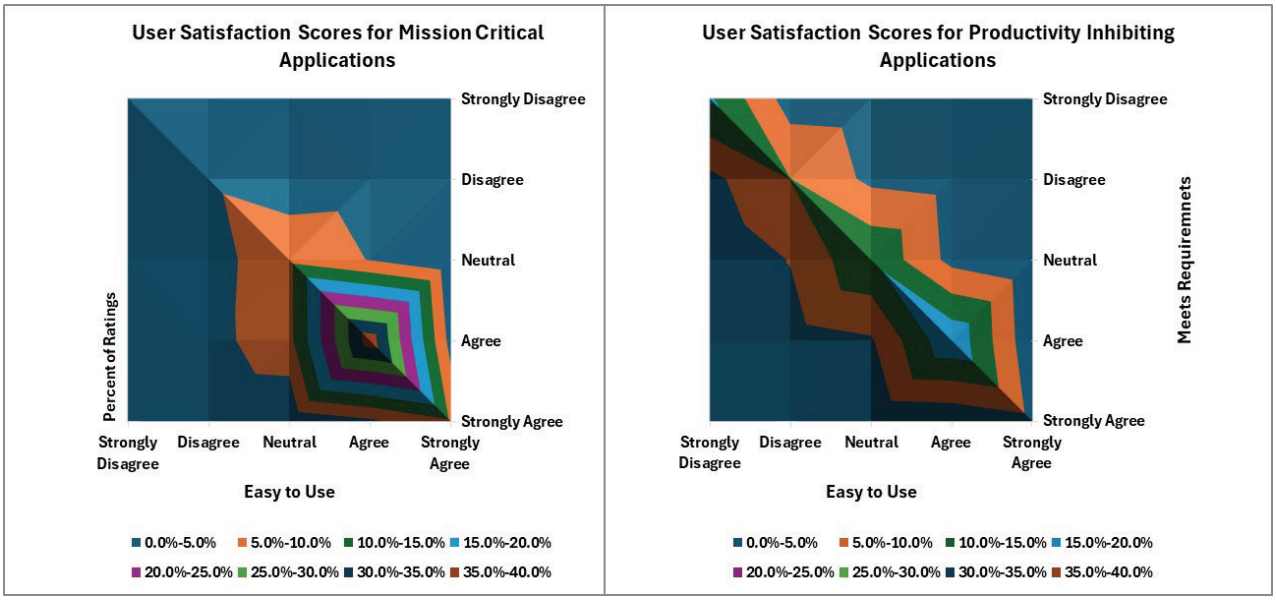
[21] Aaron Bangor, Philip Kortim, and James Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *Journal of User Experience*, Vol. 4, No. 3, May 2009.

[22] In fact, John Brooks, who is generally acknowledged as the father of software usability metrics, argues that the original Unix shell command line interface, which would likely score low on any usability index today, was highly usable in the environment for which it was created. John Brooks, "SUS: A Retrospective," *Journal of User Experience*, Vol. 8, No. 2, February 2013.

users find "acceptable" for the former might be "unacceptable" for the latter.[23] We might also expect the composition of the two concepts of "fit for purpose" and "fit for use" could differ between these two types of applications. Weapon system designers might justifiably prioritize fitness of purpose, while military pay system designers might focus more on fitness for use. For this reason, we examine these outcomes from the survey on a graph, with each measure being one of the axes of that graph, to understand how the individual elements of the scoring vary by application type. The discussions below and in Appendix B use these graphs to better understand user satisfaction.

For all mission-critical application ratings, the average score is 3.63 (out of a possible 5) for "meets requirements" and 3.56 for "easy to use," which produces a score of 65 out of 100, which is in the "marginal" range. Applications that were rated as being productivity inhibiting have an average score of 2.74 and 2.75 on these two dimensions of user satisfaction, which translates to a score of 44 out of 100, which most users find "unacceptable." Figure 4.1 provides a side-by-side comparison of the distribution of these scores across the two dimensions for mission-critical applications (left panel) and productivity-inhibiting applications (right panel). While most ratings of the mission-critical applications fall within the range of "neutral" to "strongly agree" on both dimensions (i.e., meets need and easy to use), the ratings for the productivity-inhibiting applications are more evenly distributed over the range from "strongly disagree" to "strongly agree."

Figure 4.1. Distribution of User Satisfaction Scores for All Application Ratings



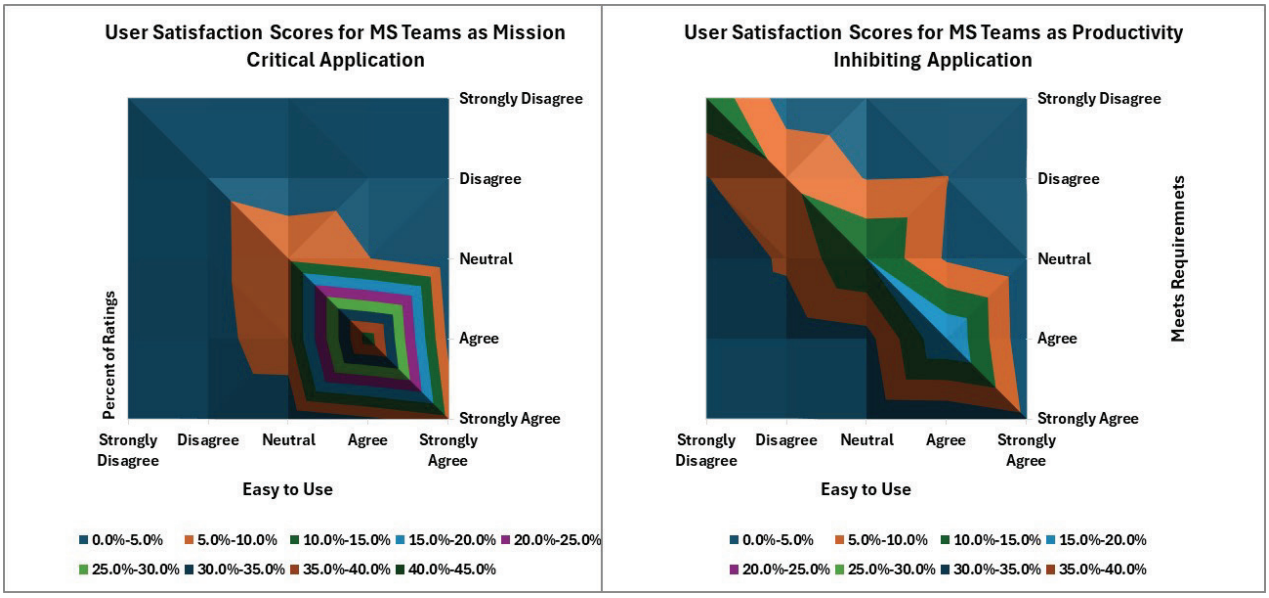
In many cases, the application being rated is the same. For instance, how users rate Microsoft Teams depends to a considerable extent on whether they see it as a mission-critical application. This is

[23] Significant research has been devoted to understanding the determinates of user acceptance of IT systems. Researchers have consistently found that, for mission critical systems, the usefulness of the system to their work is a stronger indicator of acceptance (by a factor of two) than ease of use. For an overview of these studies, see Paul Legris, John Ingham, and Pierre Collerette, "Why Do people Use Information Technology? A Critical Review of the Technology Acceptance Model," *Information & Management*, Vol. 40, No. 3, January 2003.

shown in Figure 4.2, which graphs Teams user satisfaction as a function of whether users rated it as a mission-critical application (left panel) versus a productivity-inhibiting application (right panel). For those who rated Teams as mission critical, they are relatively satisfied that it both meets their needs and is easy to use. For those who rated it as a productivity-inhibiting application, there is much less agreement that it meets their needs or that it is easy to use. In reading the open-ended comments, we found that many of those who rated Teams poorly did so because their computers are configured to automatically start the application when the computer starts up. This automated start-up consumes significant computing resources and delays those users' ability to begin meaningful work. For those on intermittent networks or in situations where the computing hardware is inadequately sized for the software needed, the automated Microsoft Teams start-up is a source of great frustration.

Therefore, we strongly recommend that local administrators review any applications that have an automatic start on computer turn-on and remove those that are not critical to the mission being performed. In particular, service personnel who share a general laptop among a dozen or more users (which both active duty and reservists tell us is common) should not have to wait for unneeded applications to start just so that they can check the current status of their orders or perform other administrative tasks.

Figure 4.2. Distribution of User Satisfaction Scores for Microsoft Teams



We close this chapter with a final caution regarding these application-centric user satisfaction scores. As with the "time lost per month" metric, we find that the application alone is rarely the sole cause of user dissatisfaction. Instead, it is more likely to be a combination of policy, network infrastructure, available computing resources, and the application that gives rise to user dissatisfaction. While the framework we describe in the next chapter advocates that technical measures and user satisfaction be used to help identify the worst-performing environments, these are indicators of issues only, not root causes.

Chapter 5

A Framework for Transparency in User Experience

DoD is an organization of millions of people undertaking multiple operations around the globe, all of which, to some extent, depend on capabilities provided by IT infrastructure and application software. A division of responsibilities, and consequently of funding, is necessary to manage the operation of this worldwide enterprise, which means that funding and technical decisions regarding IT infrastructure and application software may be made without understanding the full scope of those decisions on other organizations.

Take, for example, a technician at a CONUS Army post who uses a laptop computer to access a supply system to position an item necessary for ongoing military operations in Europe. The technician's log-in credentials are based on identity information established in the Army's directory services and the Defense Manpower Data Center and referenced by the supply system to ensure they are an authorized user. These systems, and many others involved in the transaction, are operated and maintained by different joint and Army program offices and are located in different data centers in the United States and Europe. Communications among them are managed by the Defense Information Systems Agency, Army Network Enterprise Technology Command, and other organizations, often through leased, commercial long-haul telecommunications circuits. Those organizations also provide cybersecurity oversight and monitoring of the systems involved. The need for the transaction was established by the joint task force executing the operation and conveyed to the technician through another series of systems and communications. The technician executes the supply transaction on a laptop computer provided and maintained by their unit. Long before this transaction takes place, systems required for training, duty assignment, and scheduling were employed to prepare the technician for the task.

Very few of the organizations involved in providing, managing, or maintaining the IT infrastructure and application software involved in the scenario above would even be aware they were providing a service necessary to the military operation. That awareness, and the sources and degrees of the friction added by any part of the information chain, are key to understanding and fixing IT infrastructure and application software issues that impede mission readiness. Our proposed framework, described in the remainder of this chapter, seeks to provide that awareness while making visible some of the most fundamental sources of friction.

A Framework for Assessing, Comparing, and Standardizing User Experience and Efficacy of Information Technology Systems

There is wide agreement that DoD needs a way to organize and present key information regarding the health and usability of its IT infrastructure and application software to assure that users can execute duties that require the use of those systems. To that end, NDAA 2023, Section 241, Element 3 directs "the development of a framework for assessing underperforming software and information technology, with an emphasis on foundational information technology, to standardize the measurement and comparison of programs across the Department of Defense and its component organizations." The element lists ten aspects involving the provision and operation of systems for which assessment should be possible, ranging from cost efficiency in program management, to system functionality and resilience, to user training and help-desk operations.[1] As we discuss later in this chapter, we conducted a feasibility assessment as to whether DoD currently has, or could reasonably have, a means to collect timely and accurate data needed to populate all the desired aspects listed in the NDAA language and found several items to be infeasible.

With some of the literal direction not actionable, we reconsidered the NDAA language and the results of discussions with various stakeholders and concluded that the best way to meet the evident need is to provide insight and transparency throughout the department and to Congress along three interrelated dimensions of the problem space:

- the adequacy of IT infrastructure and application software in meeting the users' needs
- the prevention of waste, encompassing both time wasted by users due to inadequate software and IT services, and resources wasted through inefficiency in the provision and sustainment of IT infrastructure and application software
- the resilience of the IT infrastructure and application software themselves, including eliminating disruption to operations caused by a failing technology base.

A useful framework exposes and makes sense of DoD data relevant to these dimensions, enabling insights that support decisionmaking throughout the department.

A Recommended New Framework Approach

A framework can be understood as a structure or system that provides a set of guidelines or principles for organizing and analyzing information or activities. To begin our research, we considered

[1] The ten aspects are

> (A) designs, interfaces, and functionality which prioritize user experience and efficacy; (B) costs due to lost productivity; (C) reliability and sustainability; (D) comparisons between—(i) outdated or outmoded information technologies, software, and applications; and (ii) modern information technologies, software, and applications; (E) overhead costs for software and information technology in the Department compared to the overhead costs for comparable software and information technology in the private sector; (F) comparison of the amounts the Department planned to expend on software and information technology services versus the amounts actually spent for such software and services; (G) the mean amount of time it takes to resolve technical problems reported by users; (H) the average rate, expressed in time, for remediating or patching weaknesses or flaws in information technologies, software, and applications; (I) workforce training time; and (J) customer satisfaction. (Public Law 117–263, 2022, Section 241, Element 3)

several well-established frameworks in the IT domain as a starting point for framework development or adoption. They include

- Capability Maturity Model Integration, a framework for improving organizational processes
- IT service management, which is focused on managing IT services effectively
- the International Organization for Standardization 9241 series (ISO 9241-201:2010), an international standard that focuses on human-centered design and the usability of interactive systems.

All three of these established frameworks have merit in that they have been matured through decades of use in government and industry; there are extensive resources available for training, certification, and assessment supporting their use in an organization; and they consider systems early in design as well as systems in operation. The scale of DoD makes direct adoption of any of them exceptionally burdensome, however, and many of the specific elements they depend on cannot be consistently measured due to the variety of program office practices and procurement methods used throughout the department. Furthermore, none is well tailored to the specific issue of lost productivity and user satisfaction. As a result, we looked for new approaches to developing a framework.

The approaches we considered were (a) those devised by the research team and labeled program-centric and portfolio-centric and (b) adoption of the Navy's Worldclass Alignment Metrics.[2] Each is discussed below. Criteria we used in considering these frameworks included

- simplicity in the information presented to decisionmakers
- ability to address both user satisfaction and user productivity
- employment of reliable metrics that enable assessment, comparison, and standardization of systems and programs with one another
- employment of metrics that have a common scale and are universally valid throughout all DoD organizations
- provision of information relating to DoD's mission, not only to programs, systems, or IT services
- leverages of industry standards
- leverages of existing DoD acquisition processes.

The program-centric approach we considered drew on the NDAA language regarding practices of DoD program management offices, which deliver IT infrastructure and software application capabilities. For this approach, we identified ten issues related to end-user experience and productivity in three broadly defined phases (requirements, solution development, and operations), along with leading and lagging indicators of whether the program effort was likely to lead to cost efficiency and a positive user experience or negative one. We based the lagging indicators on reasonably objective measures of user satisfaction through surveys and technical performance metrics measured at the end-user device—which is similar to the data we used in this report. However, this approach scores poorly on three of our criteria: the reliability of its leading indicators, which were based on subjective

[2] For completeness, we also examined an enterprise-service-centric approach similar to the ITSM. However, its metrics are based largely on profit, loss, and the user's willingness to pay, concepts that have little meaning in DoD. We did look for analogs of those concepts within DoD but ultimately rejected the approach.

information gathered from program management offices every year; its lack of integration with existing DoD acquisition processes; and perhaps most importantly, its narrow focus on a program rather than on mission. We are concerned that under a program-centric approach, issues that affect user experience and productivity that arise outside the program or system itself (e.g., network hardware failure) would be attributed to the system or program that was a victim of the problem, rather than being more properly attributed to the system, program, or non-program factor that caused the problem. As we related in our opening discussion for this chapter, merely logging in to a system via a laptop may fail for reasons outside specific systems or program elements that this approach would likely mischaracterize or obfuscate.

A portfolio-centric approach focuses on capability areas defined as sub-portfolios and portfolios of information systems. The approach we considered leverages the principles of the IT portfolio management process, which operates within the department's capability portfolio management process.[3] Thus, DoD's enterprise architecture provides a logical structure that relates information systems to the military capabilities they support and to similar systems throughout DoD. As with the program-centric approach's lagging indicators, the framework employs user-satisfaction survey data and technical performance metrics measured at the end-user device to signal the system's cost efficiency and effectiveness in meeting mission needs from the point of view of the user and end-user device. Through dynamic selection of different user groups, it also signals the source, degree, and impact of problems that arise outside of the system(s) of interest (e.g., a user group serviced by a failing network link displays poorer performance than do the users and devices served by more capable network links). In our consideration of these framework approaches, the portfolio-centric approach was the strongest in its mission focus and integration with existing DoD processes.

The Navy Worldclass Alignment Metrics framework (described in Chapter 2) has an extensive and well-structured base of metrics that roll up to present information focused for different types of interest groups, including technical support, investment management, and enterprise leadership, which potentially makes it more useful for DoD adoption. However, adopting it throughout DoD would require each organization to collect and expose a vast amount of information, beyond what is necessary for the problem at hand. The Navy, like all the other organizations charged with the responsibility to provide IT infrastructure and software to equip DoD personnel, needs sufficient information to troubleshoot technical issues and manage the enterprise in detail. In contrast, the purpose of the NDAA-directed framework is to create transparency and drive toward improvement in user experience, cost efficiency, and resilience, not to troubleshoot equipment and user support issues. Though implementing the full Worldclass Alignment Metrics framework throughout DoD would be burdensome, a subset of the operational metrics, rolled with the same logic through technology

[3] For DoD's IT portfolio management process, see DoDD 8115.01, *Information Technology Portfolio Management*, October 10, 2005; and DoD Instruction 8115.02, *Information Technology Portfolio Management Implementation*, October 30, 2006. For DoD's capability portfolio management process, see DoDD 7045.20, *Capability Portfolio Management*, September 25, 2023; and Joint Staff, Chairman of the Joint Chiefs of Staff Instruction 5123.01, *Charter of the Joint Requirements Oversight Council and Implementation of the Joint Capabilities Integration and Development System, Enclosure D*, October 30, 2021. Thus, this approach leverages existing processes within DoD, rather than inventing a new process.

outcome-driven metrics and mission outcome-driven metrics, would serve the needs of the NDAA's desired framework.[4]

In light of the strengths and weaknesses of the different approaches we considered, we recommend that DoD implement the logic and structure of the portfolio-centric framework approach, starting with a very limited dataset. Then, if the data supporting the framework are insufficient or ineffective in exposing the factors underlying poor user experience, inefficiency, and lack of resilience, the data sources should be expanded to encompass more of the Navy Worldclass Alignment Metrics. The remainder of this chapter discusses our proposed "portfolio-centric" framework in more detail.

Understanding Sources, Degrees, and Impacts of Friction in Department of Defense Software and Information Technology

The value proposition for the recommended portfolio-centric framework approach is that it exposes how differences among systems and within the information environment affect user experience and efficacy in using DoD systems. Through that transparency, this approach informs IT infrastructure and application software investment decisions and signals the source, degree, and impact of factors that impede the usability of a system. By measuring user experience at the end-user device and with the users themselves, this framework supersedes an individual system-by-system view and reports on the overall usability and usefulness of the entire chain of systems and devices that enables the user to execute their mission using IT and software.

While we describe the expected use cases for our recommended framework in detail in Appendix C, the following sections highlight specific aspects.

Combined Human and Automated Sensors Provide Actionable Information

A primary goal of our recommended framework is that it supports objective comparisons of the value of software and IT systems to the user, over time and under different conditions. As discussed in Chapter 4, *value to user* encompasses concepts of overhead time imposed on users as they use IT systems and user assessment of the usability and usefulness of a system in meeting their needs. Common examples of wasted overhead time include excessive wait times for a system to load an application or complete a common task such as sending an email. This aspect of value to the user is measured in a negative sense: More time consumed in these overhead activities lowers the value of the system to the user. Additional areas of value to the user are in help-desk responsiveness, the proactive remediation of flaws and prevention of downtime, and efficient and effective system deployment and user training.[5]

[4] We have not assessed how large a subset of the more than 300 operational metrics would need to be replicated throughout DoD, so we do not know the burden to implement compared with the value of information gained.

[5] Note that a system's value to the user may not reflect its value to the institution or its operational value. For example, a user may place little value on cybersecurity controls or mandatory training, though they are of operational and institutional value and do affect the readiness of the force to fight and perform their assigned missions.

To make implementing a framework throughout DoD manageable, we recommend that initially only two families of metrics be collected systematically from a representative sample of all users and systems. Once such a framework is operating, DoD should add another value metric on an experimental basis and assess whether that new metric enhances the output of the framework. Only metrics that significantly enhance the usefulness of the framework should be mandated across DoD.[6]

The first family is user-reported usability and usefulness of a system, measured through consistent and repeated use of the UMUX-Lite on every user-facing application. A large proportion of the department's workforce is technologically enabled to contribute to operational and institutional missions through IT and software systems provided by DoD. Consequently, the users' reports of inaccessibility or insufficiency in their IT and software tools provide essential insights into their ability to accomplish the mission. As noted in Chapter 4, the responses to the two UMUX-Lite questions can be considered separately as indications of the degree to which a named system is fit for use and fit for purpose.[7] These user-reported values signal the *mission impact* of friction arising from the system and/or the information environment.

The second family of metrics comes from measurement of technical performance, usually at the end-user device. Technical performance factors that can be captured with readily available tooling include the wait time imposed by end-user device, network, and server processing. Though these technical performance measures are associated with a named system, the measurements include contributions from all system and infrastructure factors involved (e.g., end-user device core processor loading, network communications time, server loading, and more). Note that because the data collection tools themselves create a load on end-user devices and often involve licensing costs, we recommend that they be deployed on the smallest percentage of devices necessary to produce statistically significant data.[8] The technical performance family of metrics signals the *source* and *degree* of friction caused by the system and/or the information environment.

Data from both families of metrics must include a very basic set of user demographics (e.g., assigned organization; no personally identifiable information is needed), end-user device characteristics, the geographic and/or network location of use, and date of collection event, as well as the system name to operate the framework as described below.

---

[6] For example, the Navy has started modifying help-desk contracts to report certain statistics on a regular basis in support of their Worldclass Alignment Metrics framework. Those statistics should be added to this framework data for the Navy first, and DoD should assess whether such information enhances the framework before mandating such collection throughout the department.

[7] It is important to understand that in this context we consider a user to be an agent of mission accomplishment, not an audience to be enticed, persuaded, or amused. That is, a user is a professional who carries out their duties enabled, to some extent, by DoD-provided IT and software. When that IT and software are deficient, the mission suffers.

[8] Considering potential load on an end-user device and the possibility that the device might enjoy limited use, experts from the Defense Information Systems Agency who are currently conducting experience monitoring on agency-supported computers provided the following comments. "We supplement the end user direct monitoring capability (in our case Aternity) with synthetic transaction monitoring using a non-person entity certificate in combination with our network sensors (Netscout). By implementing synthetic transaction monitoring near or in the vicinity of the end user (same building, same floor, collection of offices), we can implement 24/7 monitoring of all/most applications that any of the end users in that facility would likely use even if that use is sporadic. Additionally, the use of the non-person entity certificate gives us the capability to mimic login, send mail, file saves, opening/closing spreadsheets and any number of web transactions as if the appliance was an end user." DISA, "10/27/24 DISA Comments re Rand Report 'Underperforming Software and Information Technology in DOD,'" memo to Sarah Zabel, October 27, 2024.

Operations on Metrics Expose Source, Degree, and Impact of Problems in Department of Defense Information Technology

Because of the breadth and heterogeneity of DoD's information enterprise, different users operate software and IT systems in different information environments at any one time. Some users are in geographical locations for which the supporting network connectivity is poor, while others operate in a high-availability network environment. User experience also differs by end-user device type and age, the totality of processes running on the end-user device at any one time, the characteristics of the particular system they are operating, and other factors. Examining user experience under those different conditions provides insight into the source, degree, and impact of each of those various factors on the user's ability to use the software and IT provided to them to do their jobs. Thus, our recommended framework supports the comparison of user value factors for dynamically selected user groups that correspond to different conditions of the system(s) of interest and the information environment in which it operates.

Consider the following two notional use cases. The first is a retrospective analysis into a source of friction affecting users. The second is a prospective analysis into the potential benefit of making a particular investment.

In Case 1, a value to user rating composed of user report and/or technical performance scores for users of a particular system drops relative to the previous year. The source of the reduction in rating may come from factors within the system itself or from changes in the information environment outside the system. Potential sources of reduced user value from the system itself include a failing hardware base, rollout of a poorly designed function, or changes in the mission to which the system did not adapt. Potential sources of reduced user value from outside the system include increased network loading or failures and increased end-user device loading or failures. To investigate these potential sources, an analyst needs knowledge of when a change in mission or a rollout of a new function occurred. With that knowledge, they can examine changes in user survey responses to UMUX-Lite individually ("[Application name] meets my needs" versus "[Application name] is easy to use") and over time, covering the periods before and after the rollout of a new feature or change in mission. Separately, they can examine changes in time to complete standard transactions or tasks in the system as measured at the end-user device for user groups by geographical area, type and/or age of end-user device, or before and after the rollout of new capabilities in the information environment.

In Case 2, a military service is considering whether to replace its oldest network equipment, either with new equipment with equivalent function or in a wholesale transformation, as in the DAF's enterprise IT as a service.[9] At any one time, there are sites with users that are served by the oldest network equipment and sites with users served by the newest network equipment. Comparing technical performance scores for common software services accessed via the oldest networks to those same services accessed via newly upgraded networks provides insight into the value of the investment in terms of enhanced user efficacy.

[9] DAF's enterprise IT as a service is an initiative to upgrade IT infrastructure by purchasing both the IT equipment and the operations and maintenance of that equipment from "as a service" providers. Unfortunately, as we describe in the feasibility of gathering cost data in Appendix C, DAF is encountering difficulty in assessing the monetary return on its investment and in comparing proposed investment with current costs of operating and maintaining its IT infrastructure. Jon Harper, "Air Force Grappling with Budgetary Implications of Enterprise IT as a Service," *DefenseScoop*, July 21, 2023.

The Logical Structure of Information Technology Portfolio Management Allows Loose Association of Systems with Mission

The overarching problem that our study and framework addresses is that underperforming IT and software-based systems negatively affect DoD mission readiness and the well-being of the institution. Currently, DoD does not have an effective way to associate a user's struggles to use the systems they are provided with an operational impact or the monetary cost of any associated waste. Our framework provides the beginning of such an association by signaling the source, degree, and mission impact of issues originating within named systems and/or the information environment, under the somewhat loose association of system-to-mission provided by DoD's IT portfolio management system.

As we noted in our introductory description of the portfolio-centric framework, IT portfolio management is a component of capability portfolio management, which is described in and responsive to DoD directives for management of IT. The purpose of capability portfolio management is to "align the investments, requirements, interoperability, designs, and acquisitions of related capabilities across the DoD via enterprise portfolios to optimize operational mission capabilities across operating domains."[10] Through capability portfolio management, DoD and component organizations monitor the health of capability areas and inform improvements necessary to meet the objectives defined in the National Defense Strategy and its supporting plans, through the lens of joint, integrated mission effects. IT portfolio management is nested within capability portfolio management to address military capabilities that are implemented in software and IT. The stated purpose of IT portfolio management is to "ensure IT investments support the Department's vision, mission, and goals; ensure efficient and effective delivery of capabilities to the warfighter; and maximize return on investment to the Enterprise."[11] This framework lends transparency into IT-related factors affecting capability delivery using information from human and technical sensors.

The processes of IT portfolio management divide systems into four mission areas: warfighting, business, the DoD portion of the intelligence enterprise, and the DoD enterprise information environment. These portfolios are further divided into sub-portfolios that represent common collections of related or highly dependent information capabilities and services. Logical relationships of each system contained within a mission area are defined in taxonomies specific to those mission areas in the DoD enterprise architecture.[12] Thus, an IT system in the Army that performs a particular personnel support function is formally related to all other Army personnel support systems and to systems performing a similar function in the other services and in each DoD agency. Further, every IT system that supports a military capability formally contributes to the cost and operational success or failure of that capability. Our proposed framework brings the user's perspective into IT portfolio management and capability portfolio management, contributing to assessments of the health of each capability area and providing insight into the potential and actual effects of investments in the IT systems supporting those capabilities.

---

[10] DoD Directive 7045.20, 2023, section 3.1, p. 14.

[11] DoD Directive 8115.01, *Information Technology Portfolio Management*, pg. 2, October 20, 2005.

[12] DoD Chief, *Information Enterprise Architecture v3.0 Increment 2 Overview Document*, July 22, 2023; not available to the general public.

Implementation Across the Department of Defense Requires Governance

The proposed framework requires two levels of governance: technical and strategic. Technical governance involves tuning parameters to get the most meaning out of the data. Parameters that can be tuned include weights assigned to each element as they are combined into composite parameters, the frequency of presenting the UMUX-Lite survey for any system or application to a user, the number and type of system activities on which to collect technical performance data, and more, as described in Appendix C. The technical governing body should also recommend changes to end-user device technical performance data collection coverage (e.g., change coverage from 20 percent of all end-user devices to 30 percent) and any additional elements or families of data to bring into the framework. Technical governance should be performed by individuals from DoD and component organizations who best understand the data incorporated in the framework.

Strategic governance involves decisions that drive cost to the DoD and its component organizations and should be performed by individuals who are authorized to incur costs for their organization and best understand the mission needs served. Strategic guidance includes the required degree of coverage of end-user devices in collecting technical performance data, whether data from additional areas of "value to user" should be added to the framework (e.g., help-desk statistics), and anything else that requires dedicated action or expenditures by DoD and its component organizations. The overall purpose of strategic governance is to ensure that the framework accurately and efficiently signals the degree to which specific systems, applications, and factors of the IT environment affect the users' ability to accomplish their mission.

A Phased Implementation Across the Department of Defense Is Recommended

Our proposed framework benefits from the experience and data gained through large-scale DAF, Navy, and Joint Service Provider efforts, but there will be much to learn as it is rolled out across all of DoD. This document describes a Phase 1, baseline implementation. Growth and refinement of the framework should be centrally managed, with participation by all DoD component organizations.

Data collected by the Navy, DAF, and Joint Service Provider show that many problems affect the performance of IT infrastructure and application software. These problems are believed to arise from a combination of end-user device, network, and central system issues, including system/interface design, resiliency, capacity, and availability. Though users throughout DoD experience these problems, they are not adequately or consistently measured or characterized across the department, which impedes progress in finding solutions. Because systems used in one DoD organization may be provided by or rely on systems developed and maintained by a different organization, it is important that some system performance information be visible to all organizations.

To enable standardization and comparisons of system performance across DoD, the data underlying the framework must be consistently implemented by all organizations. However, data collection imposes burdens of cost and effort and can itself impede efficacy and user satisfaction. With that in mind, we recommend that DoD proceed with a Phase 1 implementation that requires collecting the minimum information set at the lowest sampling rate that can faithfully signal the source, degree, and impact of problems and inefficiencies. If the Phase 1 data and/or sampling rate is insufficient to

meet the objectives of the framework, additional data elements should be deliberately added to enhance the output until it is sufficient. In the same manner, the end-user device sampling rate can be increased, and individual elements of the data re-weighted as necessary to improve the results.

Phase 1 implementation of the framework employs two value measures in addition to one monetary measure. The monetary measure, yearly budgeted funding for the systems of interest in a portfolio, is currently part of the yearly appropriations and captures a portion of the cost of IT and software. We caution, however, that appropriations are made by program element, which may include one or more systems and that only some of those systems are software intensive. It is also worth repeating that no single appropriation funds all, or even the majority, of IT infrastructure and application software. Therefore, lower-level accounting records will be needed to determine the funding of software-intensive systems, which then can be combined to estimate a total cost of IT and software.[13] The value measures are not currently being collected consistently across the department. The measures we recommend be collected in Phase 1 are the components of the UMUX-Lite scale and key technical performance metrics described more fully in Appendix C.

We expect that DoD would begin implementation of our recommended framework with the DAF and DoN, asking them to make the relatively small alterations in their current data collection to meet the needs of the framework. The other DoD organizations would follow, with a phase-in period established by the department.

Limitations of This Framework Approach

We designed the framework to require the minimal initial implementation that would be effective in signaling the source, degree, and impact of problems that affect users, as well as the ability to project and analyze the impact of investments on the users. Though those capabilities address many of the framework capabilities specified in Element 3 of Section 241 of the NDAA, there are several that could be implemented in a future phase and two others for which the data do not exist in DoD. The recommended disposition of each of these sub-elements is included in Table 5.1.

Another limitation of this framework approach is that because it relies on information generated through user interactions with systems, it does not address systems in development.[14] Additionally, though DAF and DoN have both initiated end-user device monitoring, we were unable to obtain enough data to validate the framework operation. Finally, successful operation of the framework will depend on a sufficient number of users responding to UMUX-Lite surveys on a continuing basis.

[13] This is similar to the methodology RAND used to develop an estimate of the cost of DoD software for the purpose of characterizing the types of software developed by DoD. Although this process provided a valid *relative* comparison of costs at a macro level, we caution that it is likely to be highly inaccurate for absolute cost at the system level. We have yet to assess whether it will be valid at the portfolio and service levels. Triezenberg et al., 2020.

[14] A corresponding advantage is that the framework supports experimentation to improve user experience by providing a broadly instrumented environment for comparison.

Table 5.1. Recommended Disposition of Element 3 of Section 241 Specifications

Sub-Element	Specification	Recommendation
A	Designs, interfaces, and functionality, which prioritize user experience and efficacy	Incorporated in Phase 1 framework
B	Costs due to lost productivity	Can be roughly modeled or estimated in Phase 1 framework
C	Reliability and sustainability	Incorporated in Phase 1 framework
D	Comparisons between (1) outdated or outmoded IT, software, and applications and (2) modern IT, software, and applications	Incorporated in Phase 1 framework
E	Overhead costs for software and IT in DoD compared with overhead costs for comparable software and IT in the private sector	Cannot be included in framework—data do not exist in DoD. Private-sector costs for IT are typically not reported.[a]
F	Comparisons of the amounts DoD planned to expend on software and IT services with the amount actually spent for such software and services	Cannot be included in framework—data do not exist in DoD at the time when decisions must be made. Issues include the following: • Funds for software development and IT services are commonly expended through contracts and government purchase cards. • For contract funding, although funds must be obligated onto a contract before the funds' expiration date, spending may occur throughout the period of performance of multiyear contracts. As a result, true accounting of the amount spent does not exist at the time when the framework is needed. • For funds expended through a government purchase card or similar instrument, spending must be manually tracked back to the expense area (for example, software licenses or laptop computers) after the fact; this is a follow-up activity that may not be accomplished.
G	The mean amount of time it takes to resolve technical problems reported by users	Can be incorporated in Phase 2 framework but may require modification of existing help-desk and IT support contracts to collect needed data. Considering the variety of help-desk metrics in common use, it is highly unlikely that even a majority of system help desks would be collecting and reporting the same information on the same basis. For this element and the one in H below, the governance function overseeing framework implementation would need to determine the most responsive help-desk metrics to incorporate.
H	The average rate, expressed in time, for remediating or patching weaknesses or flaws in IT, software, and applications	See note in recommendation G, above.

Sub-Element	Specification	Recommendation
I	Workforce training time	Can be incorporated in Phase 2 framework, but training processes will need to be modified to collect needed data.
J	Customer satisfaction (i.e., user satisfaction)	Incorporated in Phase 1 framework implementation.

NOTE: DoD = Department of Defense; IT = information technology.

[a] In 2022, RAND researchers assessed that operations and maintenance costs for IT-intensive private-sector data and operations centers averaged approximately $2,000 per square foot of floor space. We caution, however, that these costs include more than IT and software. Bonnie L. Triezenberg, Mary Lee, Kristen Van Abel, Arianne Collopy, Brian Dolan, Sandra Kay Evans, Marissa Herron, Joshua Steele, *Essential Utilities, Developing an Investment Strategy for U.S. Space Force Mission Enabling Infrastructure*, RR-A1731-1, 2024; not available to the general public.

Chapter 6

Findings and Recommendations

Previous studies (including those by the Defense Business Board, GAO, and others) identified general issues with respect to adequacy of funding, need for monitoring, and need for governance structures to manage DoD's IT infrastructure and applications software. Our research broadly supports the findings in those studies. The accumulated results of our interviews, survey of U.S. military and civilian service members, literature review, and analysis of reports from other government organizations within and outside DoD lead to four key findings and eight recommendations, shown in Table 6.1 and discussed in more detail below.

Table 6.1. Key Findings and Recommendations

Finding	Associated Recommendation
Service members and civilians experience a variety of technical issues in using DoD-provided IT and software, some of which significantly affect productivity and morale.	Improve service and reliability for outside the continental United States secret-level internet protocol router networks.
	Regard virtual private networks as critical infrastructure and ensure appropriate redundancy and resilience.
	Conduct periodic reviews of the standard configuration and create scaled-down configurations that provide better performance to specific user types.
	Create a reliable pipeline for timely refresh of end-user devices.
Conditions throughout the service delivery chain contribute to these technical issues.	Provide mission owners and service/capability providers throughout the DoD visibility into the sources, degrees, and impacts of IT issues affecting their workforce.
	Use automated collection of IT performance data to identify the bottom 10 percent of computing environments.
The combination of authorities, resources, and responsibilities involved make the problems difficult to track and resolve.	Explore additional ways to identify and resolve IT and software problems as mission or capability issues, working beyond the traditional layered help-desk structure.
There are significant discrepancies in perceived mission impact of user issues between the users themselves and those responsible for providing the capability or service.	Strengthen the ability of mission owners and commanders to identify and address technological problems that affect mission accomplishment.

NOTE: DoD = Department of Defense; IT = information technology.

Technical Issues

Finding: Service members and civilians experience a variety of technical issues in using their DoD-provided IT and software, some of which significantly affect productivity and morale.

Various studies conducted by the Defense Business Board, DoD CIO, GAO, and individual services in the last three years consistently reflect low user satisfaction with their DoD-provided IT and software applications. In addition to these studies, described in Chapter 2, our research quantified the following impacts to productivity, mission readiness, and retention.

- DoD on-site networks are less reliable than the commercial internet used for remote access. Secret-level internet protocol router circuits OCONUS are the least performant. Ten percent of OCONUS secret-level internet protocol router users report that it takes longer than 15 minutes to gain access and begin meaningful work. This is unacceptably long and, in affecting 10 percent of the OCONUS force, is likely to affect mission readiness.
- *Average* times to log on, start applications, or recover from issues are not unreasonably long. However, ~50 percent of applications experience two or more serious issues/outages per month, significantly underperforming industry benchmarks for software application reliability. Ten to 25 percent of users experience unacceptable delays and frequent need to retry log-ons, restart applications, and recover work.
- A conservative lower-bound estimate of the cost to DoD of lost productivity due to IT and software issues for FY23 is $2.5 billion.
- While the average productivity loss when using a software application rated as "critical to mission" is two hours per month, these users experience
  - more than eight hours of productivity loss per month in one out of ten uses
  - more than 40 hours of productivity loss per month in one out of 100 uses.

  Were these outages incorporated into system availability performance metrics, we suspect they would be the dominant factor.
- Annoyance and frustration are the dominant emotions found among survey respondents. Fifteen percent of respondents exhibit signs of emotional exhaustion that could lead to retention issues. After adjusting for self-selection bias, we conservatively estimate that 5 percent of the DoD workforce may be strongly motivated to depart from service due to poorly performing IT and software.
- For users whose duties require use of a shared computer for only a few minutes per day, the average time lost when logging on, starting applications, and recovering from issues exceeds their tolerance level.

We have four recommendations associated with this finding, each of which is addressed below. The first two concern limitations we found that are related to users accessing the computing environments needed to do their job. The third recommendation involves cases in which standard software configurations adversely affect specific categories of users. The fourth addresses the need to keep end-user device hardware up to date.

Recommendation: Improve service and reliability for OCONUS secret-level internet protocol router networks.

Service through OCONUS secret-level internet protocol router networks is significantly less performant than other networks, as detailed above. Therefore, when funds for DoD network upgrades are allocated, the OCONUS secret-level internet protocol router networks should be prioritized.

Recommendation: Regard VPNs or follow-on technical solutions as critical infrastructure and ensure appropriate redundancy and resilience.

With the increase in remote access, availability of reliable VPN services equals the mission criticality of the systems accessed. Our research indicates that insufficient attention has been paid to this part of DoD's communications infrastructure to this point and that sudden and widespread service losses are common.

Recommendation: Conduct periodic reviews of the standard configuration and create scaled-down configurations that provide better performance to specific user types.

Though strict standardization of hardware and software reduces cost and increases security, in some situations, groups of users suffer unnecessary negative impacts. For example, computers that are shared by a large number of people or dedicated to limited purposes can be slower and less responsive due to switching among multiple user profiles and data. Many military functions involve little office work, though service members must occasionally complete common tasks such as training or timekeeping. Reservists also may need to focus on such tasks during limited office time on drill weekends. Systems that are unnecessarily overloaded with data that are not required for those applications can slow the entire unit's progress. Specific user types who would benefit from standardized but scaled-down configurations include:

- users of shared laptops and desktops who need quick access to a minimal set of applications
- users in mission-critical environments that require near-real time responsivity and highly available and reliable systems.

Recommendation: Create a reliable pipeline for timely refresh of end-user devices.

Though user issues arose throughout the IT and software service delivery chain, respondents to our survey identified aged, overloaded laptops as a particularly influential cause of those issues. Earlier studies referenced in this report document the average age of laptops and desktop computers for meaningful subsets of DoD. Our research delved into the impact of end-user devices in the environments in which they are operated. The combination of aged hardware with standardized software loads, centralized cybersecurity monitoring, and resource-hungry systems specialized for sophisticated functions renders the less-capable devices ineffective. While upgrading end-user devices will not solve all issues concerning DoD IT infrastructure and software, removing aged hardware as a contributor will greatly simplify interconnected and related issues.

Laptops and similar end-user equipment are typically purchased using funds that are dispersed throughout the services and agencies, easily redirected, and subject to repeated budget cuts. As an alternative, DoD should centrally plan and resource the periodic replacement of end-user devices. Data calls supporting recent Microsoft Windows operating system updates provide ample

information about where laptops and similar devices exist throughout DoD. DoD should blindly replace those devices on a reasonable schedule (nominally, every four years).

The cost of blanket end-user device replacement will reach into the hundreds of millions of dollars annually but will positively affect all mission areas. We have two suggestions for funding this effort:

- Seek a special appropriation by Congress and fence it off from "efficiency" cuts. The use of negotiated purchase vehicles (up to ten, in order to promote use of small businesses and protect against the loss of any one provider) for end-user devices and for common licenses will themselves be drivers of efficiency.
- Alternatively, tax programs based on their share of time users spend accessing their systems via end-user devices. Operating the framework described in Chapter 5 of this report can benchmark the time consumed by client computers for each monitored system.

Lack of Visibility into Information Technology and Software Issues

Finding: Conditions throughout the service delivery chain contribute to technical issues.

Deferred investment in IT and communications equipment is often cited as an important source of the problems affecting users, and our research substantiates that assertion. However, the conditions provoking these problems are much more complex than any IT system or component.

- Analysis of survey responses show that the difference between a user who experiences fast start-up and no critical issues with their mission-critical applications and those who experience more than eight and up to 40 hours of outage is not due to the application itself, but to the totality of the information environment the application runs in.
- The conditions that give rise to the more than eight hours of productivity loss per month are not easily identifiable and are unlikely to be found without changes to current monitoring procedures.
- Similarly, we find that the application alone is rarely the sole cause of user dissatisfaction. Instead, it is more likely to be a combination of policy, network infrastructure, available computing resources, and the application that gives rise to user dissatisfaction.[1]
- Survey responses as well as other research results highlight the age and overloading of end-user devices as significant contributors to the problems experienced by users.
- CIOs' top IT and software infrastructure concerns include the need for operational architectures and governance processes that integrate *across* programs of record. It is often the dependencies between programs that give rise to performance issues.
- Current DoD efforts to measure the effect of the above issues, while laudable, are not comprehensive or standardized. This limits their usefulness.

[1] For example, user dissatisfaction with MS Teams has to do with its auto-start capability and the inability for users to suppress that capability; at least some of the dissatisfaction with MHS Genesis has to do with an inability to print from the application; and the dissatisfaction with security-scanning software has to do with an inability to throttle the scans and leave resources available for mission-critical work.

Recommendation: Provide mission owners and service/capability providers throughout DoD visibility into the sources, degrees, and impacts of IT issues affecting their workforce.

Our research shows that IT and software problems affecting user productivity and mission readiness in DoD most often result from the combination of systems, support activities, and information environments involved in capability delivery rather than from point problems in any application or device. Organizations that provide capabilities and those that depend on IT and software to enable their workforce both need the ability to identify the sources and impacts of those issues to address them.

We have designed a framework for collecting, analyzing, and presenting information to help decisionmakers throughout DoD prioritize and direct resources to address problems affecting the portion of their mission accomplished through IT and software applications. With this visibility, decisionmakers can use existing DoD processes to identify and remediate gaps in technology and training to improve mission capabilities.[2] We recommend implementing Phase 1 of this framework and maturing it through the governance processes outlined in Chapter 5.

Recommendation: Use automated collections of IT performance data that emphasize finding the worst 10 percent of computing environments.

Given our finding that the 10 percent of users with truly unacceptable performance are not easily identifiable, an automated means of collecting performance metrics is needed. To be effective, this automation cannot be confined to specific networks, applications, or organizations, but must be designed to be as broadly inclusive as possible. Collecting a small number of measurements over a wide range of environments is preferrable to collecting many measurements over a more limited range of environments.

Lack of Agency and Inability to Have Impact

Finding: The combination of authorities, resources, and responsibilities involved make the problems difficult to track and resolve.

Interviews with military service CIOs revealed that policy, governance, and authority challenges are connected and pervasive, making it very difficult for the CIOs to achieve much-needed situational awareness of the scope, funding, and spending related to DoD's IT infrastructure and application software. In addition,

- Organizational culture drives some IT and software challenges and makes others difficult to overcome, adding to authority and responsibility complexities.
- Policy and governance issues affect various workforce challenges—including hiring, retention, skills development, and training—and contribute to issues in problem resolution.
- Budget challenges in how funds are allocated for IT and software and the processes for making spending decisions further complicate tracking and resolving issues.

[2] Though we intend that the information organized through the framework developed in this study be presented to mission owners and service capability providers, a reviewer commented that there may be merit in making this same information available to the users themselves.

- Current IT infrastructure and application software often fail to meet user needs due to slow implementation, lack of user involvement in acquisition, and a risk-averse culture; this underscores the importance of incorporating user needs into agile system design to better address these challenges.
- Systems and spending should align to build a cohesive operational architecture, ensuring dependable IT infrastructure and real-time data flow, with improved governance of IT and software spending to streamline problem resolution.

The overlapping spheres of responsibility inherent in the provision of information systems for an organization as large and complex as DoD leaves the user as the integrator of their own IT support. Yet, in our survey, 66 percent of respondents who included comments indicated that they felt a lack of agency—the inability to affect the issues they experienced. The user must assess if the problem they experience should be directed to a local IT support organization or a help desk specifically supporting the system they want to use. Those support resources have to determine if the problem can be resolved through their actions or if they depend on conditions in the broader information environment, including long-distance communications and cybersecurity capabilities. None of those support resources can direct all the others to take action to resolve the user's needs. No one below the Secretary of Defense has full responsibility for end-to-end system usability or effectiveness.

Recommendation: Explore additional ways to identify and resolve IT and software problems as mission or capability issues, working beyond the traditional layered help-desk structure.

Though the point of the fifth recommendation is to strengthen the chain of command with respect to technologically enabling their workforce, some situations may benefit from other actions outside established command or help-desk hierarchies. In particular, users who are served by functional systems[3] but who are under different chains of command may experience those systems differently depending on the environment provided by their local IT support organization and service capabilities. We are not trying to downplay the role of a help desk managed by a program office in supporting the use of their system or to create a second chain of command overlapping the existing chain of command. The point of this recommendation is to help users cross the different boundaries of control when necessary to resolve their IT and software issues.

In many cases, a user operates an end-user device provided and supported by a local IT organization and accesses the target system through long-distance communications provided centrally by their service and DoD, affected all the while by cybersecurity and other tools operated by other service and DoD organizations. If the chain of command is unable to identify and resolve issues related to the user's experience across all those environments, a support mechanism outside the chain of command may be useful.

We began exploration of the concept of an ombudsman-like function to address problems expressed by users of MHS Genesis, but a full treatment of the potential of that approach or of alternatives exceeded the bounds of this project (see Appendix D). DoD should assess the potential of an ombudsman or other alternative approach to identifying and resolving user issues in a manner that does not overlap or conflict with the chain of command.

[3] By *functional systems* we mean systems that provide specialized capabilities to a particular type of job series or function, such as logistics, intelligence, human resources, or financial management.

Mission Impact

Finding: There are significant discrepancies in perceived mission impact of user issues between the users themselves and those responsible for providing the capability or service.

Our research revealed discrepancies between survey respondents and IT service providers in the level of tolerance for IT and software issues they experience, with the users expressing impact to mission accomplishment to a greater degree than the service providers. The broad roll-ups and summaries of statistics regarding user satisfaction from systemwide or military department-wide studies tend to be agnostic to the mission served. The users themselves are not. When we analyze their complaints, we find that a majority of their frustration is driven by a concern that time is being wasted and that the lack of access, software capabilities, and interoperability is harming mission readiness.

Software industry standard metrics for user satisfaction assess both fitness for purpose, which speaks to military readiness, and fitness for use, which speaks to both readiness and productivity. Our survey indicates that approximately 10 percent of mission-critical application users experience eight or more hours per month of productivity loss due to IT infrastructure and application software issues. A loss of eight hours per month has significant impact on readiness, and survey respondents were vocal in explaining those impacts. In contrast to the survey respondents, however, service providers in various roles and levels whom we interviewed in a not-for-direct-attribution setting for this study indicated that "dissatisfaction" levels of 20 percent are typical and even expected.

There is some truth that a small portion of users will always find something negative to say when polled about their IT experience. However, we find very few survey respondents complained about trivia. Survey respondents were more likely (by 60 percent versus 20 percent) to refer to desired or undesired mission outcomes rather than to voice vague complaints. For example,

- Health care professionals criticizing medical IT capabilities were more likely to refer to impact on patient care than impact on themselves.
- Reservists did not merely lament that IT and software were bad; instead, they rightly pointed out the impact that time lost has on their ability to train.
- Similarly, users did not simply complain vaguely that their access to websites was blocked; they identified the websites and told us why they were needed for their mission.

Finally, it is important to note that the log-on, application start-up, recovery times, and—perhaps most importantly—the number of applications experiencing multiple serious issues per month all exceed norms the software application development industry strives for. It is not the users' imagination—DoD IT infrastructure and software do underperform when compared with industry norms.

Recommendation: Strengthen the ability of mission owners and commanders to identify and address technological problems affecting mission accomplishment.

A significant part of DoD operational success depends on people using IT and software systems, but law and DoD structure empower the institution over operations in provision and management of IT. Titles 10, 40, and 44 of the U.S. Code create DoD and military department CIOs with broad authorities with respect to budget oversight, ensuring security and interoperability of joint systems,

and compliance with law and regulation. All the CIOs we interviewed were aware of their duties and made every effort to carry them out faithfully. They showed deep concern over and involvement in infrastructure, workforce, and other aspects of IT and software that set the conditions for operational success. Over the years, however, cost and security concerns have resulted in the centralization and consolidation of IT and software management, removing control of IT from the field and from operational commanders. Now, though many DoD missions are highly dependent on IT and software for success, mission owners and commanders lack insight into or control over the end-to-end operational status of these systems. We are not recommending the decentralization and return of control of IT back to the field. Rather, we recommend that doctrine and organization be revised at many levels to empower mission owners to plan and act effectively in the information domain. Implementation of the framework described in Chapter 5 would provide mission owners visibility into the capability of IT and software systems to support their mission. This should be matched with a staff appropriate to interpret that information into plans and action to promote mission success.

Conflicting pressures on people who have a dual role of organizational CIO and of staff member responsible for effective execution of the mission in the information domain make it difficult to execute both sets of duties. DoD and military department CIOs guide their organizations to achieve greater capabilities, security, and efficiency in information systems provision and support. In contrast, a mission owner has relatively narrow interests compared with that of the institution and places resilience and effectiveness over efficiency. DoD organizations have mechanisms to prioritize and balance competing needs, but those needs must be recognized and expressed. Though a CIO with heavy institutional responsibilities could be perceived by the mission owner as a trusted and effective advocate of the mission's needs in the information domain, experience shows that that is often not the case.[4]

[4] For example, when nearly 400,000 furloughed civilians were recalled to service during the 2013 government shutdown, the Secretary of Defense's legal review determined that "Information Technology functions" were "activities that contribute to capabilities and sustaining force readiness and that, if interrupted, would affect service members' ability to conduct assigned missions in the future," but "CIO functions" did not affect the ability of people to conduct their assigned missions. Chuck Hagel, Secretary of Defense, "Guidance for Implementation of Pay Our Military Act," Memorandum for Components and Defense Agencies, Department of Defense, October 5, 2013.

Appendix A

Chief Information Officer Interview Methodology

Interview Guide

Here, we provide the generic interview guide used for semistructured interviews (Table A.1). Where relevant, we tailored language in the guide to the interview participants (e.g., for organization name or to ask about service-specific characteristics or considerations). Because we used a semistructured interview approach, we did not ask every question in the guide, and we exercised discretion with respect to follow-up questions and question order. We recorded and transcribed interviews using Microsoft Teams. These transcripts formed the data source for our qualitative analysis, described next.

Table A.1. Chief Information Officer Interview Protocol

Question	Prompt
1a. What is your position and how long have you been in that role?	• On a scale of 1 to 5, with 5 being well established and 1 being brand new, how would you rate your experience in this position?
1b. Can you describe your typical responsibilities, both day to day as well as longer term?	• What are your routine tasks? • Whom do you interact with? Supervise, report to?
1c. In your position, what would you say is your ultimate goal?	• Short-term versus long-term goals?
1d. What best prepared you to assume this role?	• What education or experiences have been the most valuable in preparing you for this role? • What does [*your service*] look for in someone to serve in a position like yours? • Is there anything you wish [*your service*] had provided you to best serve in this role?
2a. What types of decisions do you make related to software and IT issues in [*your service*]?	• What is an example of a routine/minor decision compared to a singular, more strategic type of decision? • How often do you make these types of decisions? • What drives the need for a decision (e.g., issues arise unexpectedly, routine maintenance and upgrades, policy)?
2b. How do you make those decisions?	• What factors drive your decisions (e.g., cost, time, security, usability, budget, and policy requirements)? • What information helps you make decisions? How is that information gathered? • What are the constraints that make decisions challenging? • What are the enablers that make decisions easier?

Question	Prompt
2c. Who else in [*your service*] is responsible for the types of decisions you've described?	• What other offices are involved with your ability to deliver? • How do you work with them? • In what instances have you been able/not able to take an initiative all the way to resolution?
2d. How would you describe your relationship with the acquisition community?	
2e. Once you have made a decision, what happens next?	• Whom do you share this decision with and why? • What are the protocols, if any, and how do they differ by decision type?
2f. How do you know whether the decision that was chosen led to the desired outcome for [*your service*]?	• What type of effect do you hope the decision to have on [*your service*]? • What feedback do you receive after a decision has been enacted? Have you developed any mechanisms to gather feedback? • How do you evaluate whether a decision was successful or met the desired outcome? • What is the finality of decisions? Can you explore more than one path?
3a. Can you tell us about the decisions that went into the status quo of software and IT in [*your service*]?	• What were the considerations that led to today's status quo (e.g., service needs, circumstances at the time, requirements to meet)? • In what way were you or others limited in those decisions? • What, if any, competing factors were considered (e.g., security vs. usability)? • How do legacy systems impact software and IT today? • What risks were considered when current software and IT decisions were made (e.g., vulnerabilities from delayed implementation)? • How concerned was [*your service*] about those risks? • How does risk tolerance influence decisions?
3b. How were the software and IT needs of [*your service*] determined?	• Who or what was consulted to understand needs? • What and how were the needs identified? • What capabilities were considered to address these needs? • In your role, how do you stay in touch with the ongoing needs of [*your service*]?
3c. How well does the status quo of software and IT meet the needs of [*your service*]?	• Are there specific issues or situations around user experience that you are aware of? For each, how might that issue be addressed? • Who is involved in the software/IT requirements process and how/if users are involved in vetting requirements to ensure they meet needs? • What military objectives do the current software and IT capabilities support? • How are software and IT capabilities enabling or hindering these objectives? How does this differ across the different organizations in [*your service*]? • If software and IT capabilities are insufficient for meeting objectives, what are the consequences?
3d. What are the costs associated with the status quo of software and IT in [*your service*]?	• From what color of money do these funds originate? • Who receives funds to develop or maintain the status quo software and IT?

Question	Prompt
	• What other software- and IT-related costs does [*your service*] incur? • Does the availability of funds meet the demands of status quo software and IT expenses?
3e. What characterizes the IT and software issues in the department?	• What are the most pressing IT and software issues in the department? • How do these pressing issues vary across (a) functional areas and (b) CONUS vs. OCONUS locations?
3f. Are there any policy challenges at play?	• What are they?
4a. How could the current state of software and IT be improved?	• What are the short-term and long-term potential improvements? • How would improvements impact the risks we discussed earlier? • Would these improvements introduce new risks, and if so, what are they? • How is interoperability considered? • How is usability considered?
4b. What would be needed to make these improvements?	• What is the understanding of the capabilities needed to achieve these improvements? • What are some of the greatest hurdles to overcome for achieving these improvements? • What is the understanding of the required costs relative to budget?
4c. How does the state of the workforce enable these improvements or prevent them from taking place?	• To what extent does the workforce meet the skills, experience, and availability that is needed? • How are workforce needs met internally versus externally (e.g., internally grown talent vs. industry hires)? • What is the role of contractors for improving software and IT, and what are the associated benefits and limitations to working with industry?
4d. What would the implementation of improvements look like?	• What are the timelines for making improvements, and how do they compare with the expressed need/demand? • What supports implementing improvements? • What barriers are there to implementing improvements?

Qualitative Analysis Methodology

Once all interviews were completed, we conducted a thematic analysis of the interview data.[1] We developed a codebook using a data-driven approach to analyze thematic patterns across the interview notes.[2] Table A.2 provides definitions for all codes.

[1] R. E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*, Sage Publications, 1998; Matthew B Miles, A. Michael Huberman, and Johnny Saldaña, *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed., Sage Publications, 2014.

[2] In thematic coding, codes are "meaningful labels" that we apply to parts of text documents to analyze qualitative data. See Jessica T. DeCuir-Gunby, Patricia L. Marshall, and Allison W. McCulloch, "Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project," *Field Methods*, Vol. 23, No. 2, 2011, p. 137. In using a data-driven approach to codebook development, we took an inductive approach to developing codes with the goal of

Two researchers coded three of the same transcripts and then discussed the codes and coding procedures at length in order to refine the codebook—including definitions, examples, and inclusion and exclusion criteria—and to reach a consensus about how codes should typically be applied. For some codes, we included examples of what could be included or excluded to help further refine our code definitions. After two rounds of practice coding and discussion for the purpose of codebook refinement, two researchers split the coding (i.e., each coded three interviews in Dedoose, a program designed for qualitative and mixed methods analysis).[3] We held follow-on discussions to assess how codes were applied and used the qualitatively coded data to examine patterns and themes across topic areas.

Table A.2. Thematic Codebook Used for Chief Information Officer Interviews

Code	Definition
Challenge	An identified challenge related to software/IT
Enabler	An identified enabler related to software/IT; may relate to implemented solutions
Authority	Authority to make decisions that affect service IT/software
Policy	Reference to specific or general policies related to IT/software
Governance	The processes, structures, and systems that are used to manage and make decisions about IT/software
Partnerships	Relationships and interactions with others external to the service, including other services and industry
Standards	Standard approaches/models that can be used DoD-wide
Infrastructure	State of service buildings, permanent installations, and equipment as it pertains to IT/software, cloud
Commercial technology	Use of commercial technology as it relates to service IT/software
Network	Comprehensive system of information capabilities and processes that are interconnected and end-to-end. The system is designed to collect, process, store, disseminate, and manage information on-demand for warfighters, policymakers, and support personnel
Data management and digital integration	Data storage, centralization vs. stovepipes, cloud use, sharing capabilities, DevSecOps, technical stack
Interoperability	Ability to operate between systems originating in different functions, services, and so on
Cybersecurity	The capability of allies, whether within or across services, to work together in a coordinated, effective, and efficient manner to achieve tactical, operational, and strategic cyber objectives
Last mile	Delivery of IT/software services to customers located at their area of operation

reducing our raw interview notes into smaller thematic categories, as compared with other approaches such as theory-driven coding, which focus on a priori codes. See DeCuir-Gunby, Marshall, and McCulloch, 2011.

[3] Dedoose version 9.0.17 (2021) is a web application for managing, analyzing, and presenting qualitative and mixed method research data developed by SocioCultural Research Consultants, LLC.

Code	Definition
Divesting	Removing of systems/processes/IT/software that is not wanted or needed any longer; unnecessary complexity due to not divesting
Redundancy	Redundant efforts, spending, and the like due to lack of alignment/integration
Sustainment	Long-term sustainment of new capabilities/abandonment due to lack of sustainment
Budget process	How the budget process supports or hinders IT/software needs
Budget sufficiency	The sufficiency of budget funds to meet IT/software needs
Identifying user needs/ requirements	Efforts and process in place to identify user needs or set requirements
Meeting user needs	Extent to which service is meeting user needs; employee satisfaction
Implementation	Process to implement IT/software, including concept of operations
Culture	Common attitudes, beliefs, and behaviors within the service
Workforce hiring	Process and ability to hire talent
Workforce retention	Experiences retaining talent
Workforce skills	How the workforce skills are meeting demands or evolving with needs
Workforce training	Workforce training efforts
Strategic planning	Vision, future planning, strategic documents that guide IT/software efforts
Solution: implemented	Solutions that have been/are being implemented (e.g., Zero Trust, Microsoft 365)
Solution: recommended	Suggestions/recommendations for solutions

NOTE: DoD = Department of Defense; IT = information technology; DevSecOps = Development, Security, and Operations.

Appendix B

Survey Instrument and Responses

Survey Instrument

We provide the survey instrument used to measure user satisfaction and experience below. Instructions for the survey programmer are contained within brackets << >>.

<<Opening screen>>

Informed Consent: The RAND Corporation has been contracted to conduct the NDAA 2023 Section 241 study as an independent nonprofit institution. NDAA 2023 Section 241, "Study on Costs Associated with Underperforming Software and Information Technology," requires DoD to conduct an independent study on the challenges, effects, and potential solutions related to the use of software and information technology (IT) within DoD. This study involves a survey of Armed Forces members to measure lost working time. We will ask you about your most recent log-on experience and the performance of software needed to do your job.

This survey should take approximately 10–20 minutes. Your participation in this survey is entirely voluntary. You can choose not to participate or skip any questions that you do not feel comfortable discussing. In our reporting, we will aggregate the survey responses and will not be attributing your responses to you. Though we can never promise absolutely zero risk of being identified, we anticipate participation poses very minimal risk, given aggregating responses and controlling contact information we have for participants. Please do not discuss or comment on classified information.

<<Screen break>>

1. **Please specify whether your current duty location is:**
   - ☐ Outside of the continental United States (OCONUS)
   - ☐ Within the continental United States (CONUS)

2. **We will ask you about the performance of IT and software on the networks most critical to the performance of your duties. Please identify the networks you use to perform your assigned duties (select all that apply):**
   - ☐ NIPR
   - ☐ SIPR
   - ☐ JWICS
   - ☐ Other (please specify):
   - ☐ _____

<< Questions 3 to 7 are repeated for the first two networks checked>>

3. **Today, will you access the network identified in question 2 directly or are you a remote user?**

   ☐ Direct
   ☐ Remote
   ☐ Will not access it today

4. **At your current duty location and for the network identified in question 2, is network connectivity generally:**

   ☐ High speed and reliable
   ☐ High speed but intermittent
   ☐ Low speed but reliable
   ☐ Low speed and intermittent
   ☐ Often down for more than an hour

5. **For your most recent log-on to the network identified in question 2, how many minutes did it take from when you began the process until meaningful work could be started?**

   ☐ Less than 1 minute
   ☐ Between 1–5 minutes
   ☐ Between 5–15 minutes
   ☐ Between 15–30 minutes
   ☐ More than 30 minutes and/or I was unsuccessful.

6. **How does the duration you reported in question 5 compare to your usual experience?**

   ☐ Much better than usual
   ☐ Slightly better than usual
   ☐ About the same
   ☐ Slightly worse than usual
   ☐ Much worse than usual

7. **How many years have you worked with the network identified in question 2 or in a similar IT environment?**

   ☐ Less than 1 year
   ☐ More than 1 year

<<Screen break>>

In the next section, we ask about your experience using IT applications (or suites of applications) to perform your duties. A suite of applications might be those bundled together and accessed through a software factory, test environment, workplace as a service, etc. Applications that are embedded in a radio, sensor, weapon or vehicle are outside the scope of this survey.

8. **Name up to three IT applications that are critical to performing your assigned duties (i.e., without them it would be difficult to complete your assignment).**

9. Name up to three applications that you find unnecessarily hinder your productivity. This can include applications that are not critical to your assigned duties.

<< Separate Screen for each application listed in Q8 and Q9—respondents answer the same questions below for each application. Accept no more than 6 total applications per respondent.>>

10. Please estimate the percent of your on-duty time spent using [application name]? (Use includes passive activities, i.e., if you have a chat application open and constantly monitor it for new information, then usage is 100 percent.)

- ☐ Less than 25%
- ☐ 25–50%
- ☐ 50–75%
- ☐ 75–100%

11. For your most recent use, how many minutes did it take from when you decided to use [application name] until you could begin meaningful work?

- ☐ Almost immediately
- ☐ Less than 1 minute
- ☐ Between 1–2 minutes
- ☐ Between 2–5 minutes
- ☐ More than 5 minutes

12. How does the duration you reported above compare to your usual experience over the last month?

- ☐ Much better than usual
- ☐ Slightly better than usual
- ☐ About the same
- ☐ Slightly worse than usual
- ☐ Much worse than usual

13. Please indicate your level of agreement with the statement "[application name] meets my requirements."

- ☐ Strongly disagree
- ☐ Disagree
- ☐ Neither agree nor disagree
- ☐ Agree
- ☐ Strongly agree

14. Please indicate your level of agreement with the statement "[application name] is easy to use"

- ☐ Strongly disagree
- ☐ Disagree
- ☐ Neither agree nor disagree
- ☐ Agree
- ☐ Strongly agree

15. Over the last month, have any of the following issues occurred in your use of [application name] (choose all that apply):

- ☐ Application could not be accessed.
- ☐ Application failed to open, did not respond, or timed out.
- ☐ Application closed unexpectedly.
- ☐ Application lost previously generated or input information.
- ☐ Application had no issues.

16. For the most recent issue you had in using [application name], how many minutes did it take from when the issue occurred until you could resume meaningful work?

- ☐ Less than 5 minutes
- ☐ Between 5–15 minutes
- ☐ Between 15–30 minutes
- ☐ Between 30 minutes to 1 hour
- ☐ More than 1 hour and/or I was unsuccessful.

17. How does the duration you reported in question 16 compare to your usual experience over the last month?

- ☐ Much better than usual
- ☐ Slightly better than usual
- ☐ About the same
- ☐ Slightly worse than usual
- ☐ Much worse than usual

18. <<Ask only if response to question 15 includes loss of information>> **Please estimate the number of hours you spent over the last month recreating lost products.**

- ☐ Less than 1 hour
- ☐ Between 1–8 hours
- ☐ Between 8–20 hours
- ☐ Between 20–40 hours
- ☐ More than 40 hours

<<Screen break>>

19. Is there anything else we should know about how the use of software and information technology impacts your productivity (either negative or positive)? [200 words or less]

Survey Response Rate and Engagement

While the overall survey response rate is 8 percent, it is unevenly distributed across the services and across the military/civilian divide. Statistics are provided in Table B.1. In general, newly formed components and organizations (e.g., Space Force, Defense Health Agency) had higher response rates than the more established services. Similarly, we observed a higher response rate for civilians than for uniformed military personnel. Of the services, the Navy viewpoint is underrepresented in our results.

However, we find little evidence that this underrepresentation biases the results of our research. In general, service affiliation and military versus civilian status are uncorrelated with how personnel regard their experience using DoD's IT infrastructure and software.

Some of the differences in response rates between military and civilian may be caused by differing policies that determined whether our invitation emails were sent to spam folders or had their links removed, as evidenced by the fact that the number of reported undeliverable/blocked invitations is higher for military than civilian personnel.[1] Similarly, some of the difference in response rates is undoubtedly due to the fact that many uniformed military do not use computers to complete their assigned duties and either did not check their military email accounts in the survey window or did not have access to a computer to complete it.[2] We received several reports of local IT access control measures that blocked military personnel from accessing the on-line survey. [3] While the above undoubtedly explains part of the difference, we suspect that much of the difference in response rates is due to the Space Force and civilian personnel having more up-to-date email addresses in the Defense Manpower Data Center database we used for our sample. For the Navy and Marine Corps, we were able to identify a sizable block of uniformed military email addresses where the domain was simply invalid.[4]

For those who did respond, engagement was high. All questions could be skipped, yet 99 percent completed all questions regarding access to one of the networks they use in the conduct of their duties. Two-thirds of respondents rated at least one mission-critical application, and one-third also rated an additional application that inhibited productivity. One-third wrote replies (some quite lengthy and all of them thoughtful) to our open-ended question, "Is there anything else we should know about how the use of software and information technology impacts your productivity (either negatively or positively)?" Table B.2 shows engagement statistics.

[1] We caution, however, that for operational security reasons, some DoD email systems suppress the notification of undelivered emails. Our count of undeliverable/blocked is a lower bound; the actual count may be much higher.

[2] The survey was open for six weeks after the initial invitations were sent. We sent reminder emails twice during the open period.

[3] Highly motivated personnel who encountered blocked access told us that they then sent the survey invitation email to their home computers and accessed our links from there to complete the survey.

[4] These were @training.navy.mil addresses that we are told have never existed. In Table B.1, we count these as "invalid domain."

Table B.1. Survey Response Rates

	Air Force		Army		Marine Corps		Navy		Space Force		Health .mil	Other Civilian	Total Sample
	Military	Civilian	Military	Civilian	Military	Civilian	Military	Civilian	Military	Civilian			
Invitations sent	15,737	5,134	26,281	8,878	6,535	420	15,876	5,816	369	11	1,140	1,873	88,070
Completed	1,029	581	699	1,163	227	64	362	386	44	1	231	142	4,929
Partially completed	364	200	195	365	83	24	104	117	10	0	98	59	1,619
Refusals	0	0	0	0	0	0	0	0	0	0	0	0	0
Invalid domain	0	0	0	0	300	0	3,451	3	0	0	0	0	3,754
Undeliverable/ blocked	4	102	520	50	15	2	166	29	0	0	2	95	985
Response rate	6.5%	11.5%	2.7%	13.2%	3.6%	15.3%	3.0%	6.7%	11.9%	9.1%	20.3%	8.0%	5.9%
Partial response rate	8.9%	15.5%	3.5%	17.3%	5.0%	21.1%	3.8%	8.7%	14.6%	9.1%	28.9%	11.3%	7.9%

NOTE: Response rate = completed/(total sample size—invalid domain—undeliverable/blocked).

Table B.2. Engagement Statistics

Completed Surveys			Network Qs		Mission-Critical Applications			Problem Applications		
	Surveys Started	Surveys Completed	One Network	Two Networks	One Critical App	Two Critical Apps	Three Critical Apps	One Extra App	Two Extra Apps	Three Extra Apps
Counts	6,546	4,927	6,461	1,828	4,426	4,088	3,478	2,193	1,376	828
Percentage		75%	99%	28%	68%	62%	53%	34%	21%	13%

NOTE: Percentages are of the surveys started. All network and application ratings are included in the analysis, regardless of survey completion status.

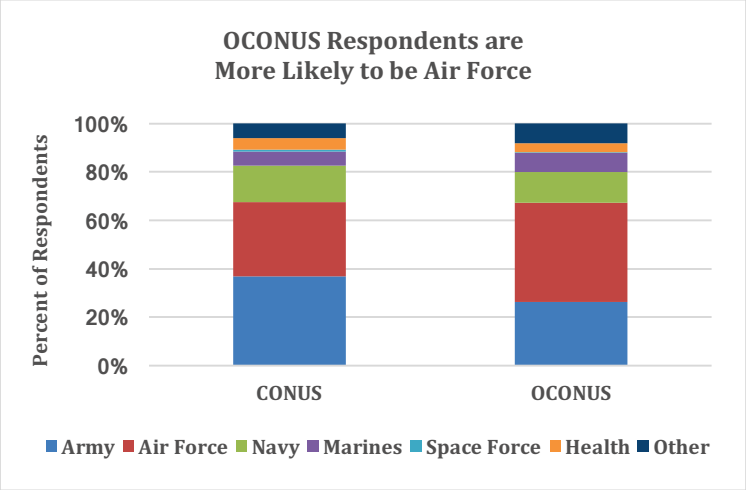
Results by Question

Here, we provide the full results of the survey, organized by question. In many cases, the text of this appendix is identical to that used in the summary provided in the main body of the report. It is repeated here within the fuller context of the analysis.

Q1. Duty Location

Fourteen percent of respondents reported that they are located OCONUS, which is representative of the total force. Of those, the largest contingent is from the Air Force, as shown in Figure B.1. While we know from the responses to Question 19 that some shipboard personnel did receive and respond to the survey, we suspect that, given the Navy's lower response rates, naval personnel deployed on ships OCONUS are underrepresented in the OCONUS sample.

Figure B.1. Service Affiliation of Respondents



NOTE: OCONUS = outside the continental United States.

Q2. Networks Used in the Performance of Duties

Seventy-five percent of respondents indicated that they use the non-classified internet protocol router network in the performance of their assigned duties, 21 percent listed the secret-level internet protocol router network, 4 percent listed the Joint World-Wide Intelligence Communication System, and 20 percent indicated that they used some other network. Note that respondents were allowed to list all networks that applied, and so these percentages do not add up to 100.

Q3. Remote Versus Direct Network Access

Almost all (96 percent) of those who listed a non-classified internet protocol router as necessary for their assigned duties logged on to it on the day of the survey, with 23 percent logging on remotely. As Table B.3 shows, the percentage of remote log-ons is lower for those who use classified networks.

Table B.3. Remote Versus Direct Network Access by Network

Network	Direct Access	Remote Access	Did Not Access Today
Non-classified internet protocol router	73%	23%	4%
Secret-level internet protocol router	65%	5%	30%
Joint World-Wide Intelligence Communication System	76%	4%	20%
Other	66%	22%	12%
All networks	71%	19%	10%

When comparing the service affiliation of the responses, we find that the Army has by far the highest proportion of remote log-on users at 30 percent. The Navy is a distant second, with 18 percent of its respondents logging on remotely on the day of the survey.[5] Only 15 percent of Air Force respondents logged on remotely, and only 13 percent of Marines and Space Force respondents did so as well.[6]

Q4. Network Ratings

Users' reported experience with these networks varies. Users of a non-classified internet protocol router generally rate it as high speed (64 percent). Regardless of how they characterize its speed, half of all users of a non-classified internet protocol router characterize it as intermittent. Secret-level internet protocol router users are more evenly distributed in their ratings, with half rating it high speed and half as low speed. As with the non-classified internet protocol router ratings, regardless of speed, half of all secret-level internet protocol router users characterize it as intermittent. The Joint World-Wide Intelligence Communication System fairs slightly better in that 69 percent of users rate it as high speed, and only 37 percent characterize it as intermittent.

We also broke out these ratings based on whether users reported accessing the network directly versus remotely and found statistically significant differences. Remote users of a non-classified internet protocol router are more likely to rate it as high speed (70 percent versus 64 percent for direct-access users) and less likely to rate it as intermittent (43 percent versus 51 percent for direct-access users). These percentages indicate that *commercial network service providers are perceived as providing better reliability than DoD's on-premise networks*. Network ratings for remote secret-level internet protocol router users are also slightly better than ratings of those who access it directly. We did not have sufficient sample size for remote users of the Joint World-Wide Intelligence Communication System to generate a valid comparison.[7]

Finally, we broke out these ratings based on whether the respondent was uniformed military versus civilian. Civilian users are much more likely to rate the non-classified internet protocol router

[5] At the time of the survey, the Navy's primary VPN had recently been removed from service. If we add in the large number of personnel who complained about the loss of the VPN, the percentage of Navy users who log on remotely would rise to 28—equivalent to the Army remote log-on rate.

[6] Due to its small size, there is an uncertainty of +/− 5 percentage on any percentages of the Space Force provided in this report.

[7] Only ten users of the Joint World-Wide Intelligence Communication System report accessing the network remotely.

network as high speed (70 percent versus 55 percent for the uniformed military) and less likely to report that it was intermittent (45 percent versus 50 percent). There was less difference for the classified networks. We suspect the better network performance experience reported by civilians is because uniformed military are more likely to access networks in more austere locations.

Network ratings for "other" networks closely mirror those reported for non-classified internet protocol routers.

Q5. System Access Times

Overall, half of all respondents reported that the time it takes from when they begin the process of logging on to DoD systems until meaningful work can be started is under one minute. Only 13 percent reported that it takes more than five minutes. However, a small subset within that group (4 percent) reported that it takes more than 15 minutes to access their systems and begin meaningful work. For many of them, this involves multiple restarts of their computers. For those users, this lengthy access time is a source of great frustration. As one user told us,

> At least three times a week, it takes between 24 minutes and 64 minutes for me to be able to start using my [DoD-supplied] laptop from the time I try to log in. I've started keeping a log because the wait time is so absurd. . . . Eighty percent of the time, my computer starts up with a black screen and no icons. I constantly have to restart my computer because it is non-responsive. . . . It is so bad, I don't see how it is not a national security issue.

These results for system access time do not significantly change by service affiliation, network, or whether the respondent is based OCONUS, which makes it difficult to find the small subset of users who suffer these long (and, to us, unacceptable) wait times.

Remote access users reported slightly longer access times, but the average additional time is less than one minute. In return for this small additional access time, users experience the higher reliability and speed of the non-DoD networks—a trade that appears to be well worthwhile. Although it is unclear whether these external networks will be able to sustain that performance when under attack, remote access appears to improve productivity for DoD during peacetime.

OCONUS users of classified networks also reported slightly longer access times, with the average additional time approaching two minutes. Ten percent of OCONUS secret-level internet protocol router users report that it takes longer than 15 minutes to gain access and begin meaningful work. This is unacceptably long and, with 10 percent of the OCONUS force experiencing such delays, these delays are likely to affect mission.

Q6. Usual Access Time

Recall that we asked users to report their most recent access time (as opposed to asking about a typical time, which can bias results). To ensure that our results are not biased by having respondents answer on a particularly bad or good day, we then followed that with a question as to whether their reported time was better or worse than their usual experience. A significant majority of respondents (83 percent) indicate that the reported time was "about the same" as usual. For the remaining

respondents, half stated the time was better, and half stated it was worse. There is no skew to the results. Therefore, we are comfortable saying that the results for access time are relatively unbiased by temporal considerations or saliency bias.

Q7. Years of Experience Using Similar Networks

New users of IT systems are more likely to rate performance harshly (because it is different from what they are used to) than those who have had time to adapt their work habits to the peculiarities of a specific system. Seven percent of respondents told us that they had less than one year of experience on a similar network or IT environment. This is a reasonable percentage for any organization and leads us to conclude that our results are not skewed by a disproportionate percentage of either new or experienced personnel.

Q8. Mission-Critical Applications

We received approximately 12,000 ratings of applications that respondents stated were critical to the performance of their duties. Of these, nearly one-third (~4,000) were for Microsoft Office Word, Excel, and PowerPoint. Another one-sixth (~2,000) were for Microsoft Teams. Another ~750 were for various browsers. Approximately 700 respondents rated Adobe's products—dominated by Acrobat, with its integrated ability to digitally sign documents using certificates on DoD-issued common access cards. MHS Genesis was the next most often rated application with ~250 ratings. Only 36 applications (or family of applications) received more than 20 ratings, and these make up two-thirds of the total number of ratings received. The last one-third of the ratings are for applications with fewer than 20 ratings. The dominance of Microsoft products, which are generally rated quite favorably by respondents, pulls up the mean scores for both productivity and user satisfaction. The one exception to this is Microsoft Teams, which, as we will discuss next, has quite unfavorable ratings from respondents who did not cite it as being critical to their assigned duties.

Q9. Productivity-Inhibiting Applications

Of the approximately 6,400 ratings received for applications that respondents considered to be productivity inhibiting (and were not among their top three mission-critical applications), roughly one in ten (~650) were for Microsoft Office Word, Excel, and PowerPoint; and one in 20 were for Microsoft Teams. These respondents were 22 percent less likely to rate Word, Excel, and PowerPoint as meeting requirements and being easy to use and 32 percent less likely to rate Teams as meeting requirements and being easy to use.[8] Another one in 20 of the applications rated as being productivity inhibitors were commonly used cybersecurity applications (automated scanning software, common access card certificate-based access controls, website blockers, and so on); this category of applications

[8] These losses were in both components of the UMUX score; that is, respondents rated these applications not just as failing to meet their needs, but also as being productivity inhibitors. For Teams, much of this dissatisfaction appears to arise from the fact that Teams is often configured to start up at log-on time. For those users who do not use Teams as part of their duties, the extra time and complexity introduced by this auto-start feature is a significant source of annoyance and frustration.

is largely nonexistent in the list of applications users cited as critical to the performance of their duties.[9] The next most rated family of applications were the various browsers, followed by Adobe products. Only 24 applications (or family of applications) received more than 20 ratings, and these make up only 40 percent of the total number of ratings received. This means that there is significant diversity in the applications that are rated as productivity inhibiting, with nearly 60 percent of the ratings being for applications that have fewer than 20 complaints; identifying and prioritizing these applications are extremely difficult.

Q10. Percentage of On-Duty Time Spent Using the Application

To better understand mission and productivity impacts of application performance issues, we asked respondents to estimate the percentage of their on-duty time that the application is in use. This includes passive use, such as the need to monitor a chat screen for input (if the screen is open for the entire on-duty time, then usage is 100 percent). Mission-critical applications tend to be used throughout the respondent's day, with 35 percent of them in use for more than three-quarters of the on-duty period and 60 percent in use for more than half of the time. If these applications are unavailable for even a few hours, there will be impact to mission.

In comparison, the applications that were named as productivity inhibiting are used for shorter durations in the respondent's day. Half of all productivity-inhibiting applications are in use for less than a quarter of the respondent's duty period, and only one-third are in use for more than half of the period. If these applications are unavailable for a few hours, it is likely that users can work around the outage.

Thus, we are comfortable making the general assumption that time lost when using mission-critical applications has a direct impact on mission as well as on productivity, while the impact of time lost when using applications respondents classified as productivity inhibiting is primarily on productivity, not mission.

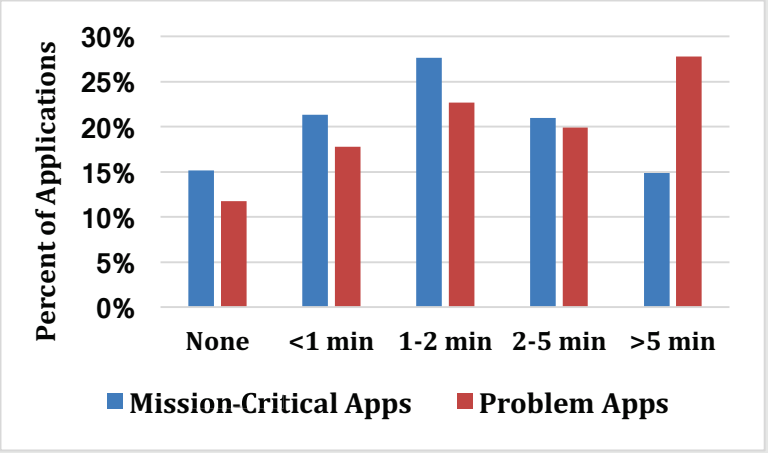
Q11. Application Start-Up Time

For each application a respondent listed, we asked them to report the length of time it took to start it. The distribution of the responses is shown in Figure B.2, segregated by mission-critical applications versus productivity-inhibiting applications (referred to colloquially as "problem apps" in the figure). Ideally, applications should start up instantaneously or in less than a minute.[10] However,

[9] Menlo security and Tanium were the most often cited by name. This is not to say these applications are particularly annoying; it could also be that their names are better known than those of other cybersecurity products.

[10] Our review of the literature on human-machine interactions indicates that a computer's response time of a tenth of a second feels instantaneous to the user. One second feels to the user as if they are operating the machine directly with no interruption in flow of thought for the user, but a delay of ten seconds interrupts users' thought processes. For more information on human-machine interactions times, we recommend the original research by Miller, 1968.

Figure B.2. Application Start-Up Time Comparison



NOTE: The large number of responses in the "more than 5 minutes" category in this figure means the distribution of responses is neither a bell curve nor uniform. A strictly algebraic computation of the mean is therefore invalid and results in a mean that is too low. Future survey designs should add a "5–10 minute" category, with the final choice being ">10 minutes."

according to respondents, over one-third of mission-critical applications and nearly one-half of problem applications take more than two minutes to start-up.[11] On average, mission-critical applications take from 1.5 to 3.5 minutes to start up, while and problem applications take from two to 4.5 minutes.[12] For those who experience multiple issues per day with an application, these start-up times can quickly become a source of frustration.

Q12. Usual Application Start-Up Times

As with the network access times, a majority of respondents (80 percent) indicated that the reported application start-up time was "about the same" as usual. For the remaining respondents, half stated the time was better, and half stated it was worse. There is no skew to the results. Therefore, we are comfortable saying that the results for application start-up time are relatively unbiased by temporal considerations.

[11] Clearly, the average application start-up times experienced by DoD users can interrupt flow of thought. This disruption may not affect mission or productivity if users start applications only at the beginning of a shift or if users expect delays and adjust to them. However, even when delays are repeatable and expected, research has found that delays that impede the necessary pace of the task at hand will result in productivity loss and user dissatisfaction. Dabrowski and Munson, 2011.

[12] The high end of our estimate for average start-up time of the productivity-inhibiting applications is undoubtedly low. The large number of responses in the "more than 5 minutes" category means the distribution of responses is neither a bell curve nor uniform. A strictly algebraic computation of the mean is therefore invalid, and results in a mean that is too low. Future survey designs should add a "5–10 minute" category, with the final choice being ">10 minutes."

Q13 & Q14. User Satisfaction

We use an industry standard measure for evaluating user satisfaction with software applications, UMUX-Lite.[13] This comprises two of the nine questions asked for each reported software application. The two questions are: "Please indicate your level of agreement with the following statements: '[application name] meets my requirements,' and '[application name] is easy to use.'" Agreement is on a five-point scale from strongly disagree to strongly agree. Scores can therefore range from 1 to 5, with 5 being best.

Table B.4 contrasts the average user satisfaction scores of the five most often rated mission-critical applications and the five most often rated productivity-inhibiting applications. Recall that we provided the opportunity to rate only three most critical applications, and we did not ask respondents to distinguish between mission-critical and non-mission-critical when citing their top three productivity-inhibiting applications.[14] Therefore, one should not assume that mission-critical applications are not productivity inhibiting or that productivity-inhibiting applications are not also mission critical (just not in the top three).

Table B.4. Most Often Rated Applications

	Mission-Critical Ratings		Productivity-Inhibiting Ratings	
Rank	Application Family	Average UMUX Score[a]	Application Family	Average UMUX Score[a]
1	MS Office Word, Excel, PowerPoint	3.7/3.8	MS Office Word, Excel, PowerPoint	3.3/3.2
2	MS Teams	3.7/3.7	MS Teams	2.7/2.8
3	Browsers	3.8/4.0	Cybersecurity applications	2.0/2.2
4	Adobe	3.7/3.6	Browsers	2.9/3.3
5	MHS Genesis	2.8/2.4	Adobe	3.2/3.2

NOTE: UMUX = Usability Metric for User Experience. Presented as meets requirement score/easy to use score.
[a]Scores are on a scale of 1 to 5, with 5 being best.

Of note in Table B.4 is the relative symmetry of the two dimensions of user satisfaction. In most cases we examined, respondents make very little differentiation in their scoring of whether an application meets their needs versus whether it is easy to use. The table shows two exceptions to this rule:

- Respondents were more likely to rate MHS Genesis as meeting their needs than they were to rate it as easy to use.

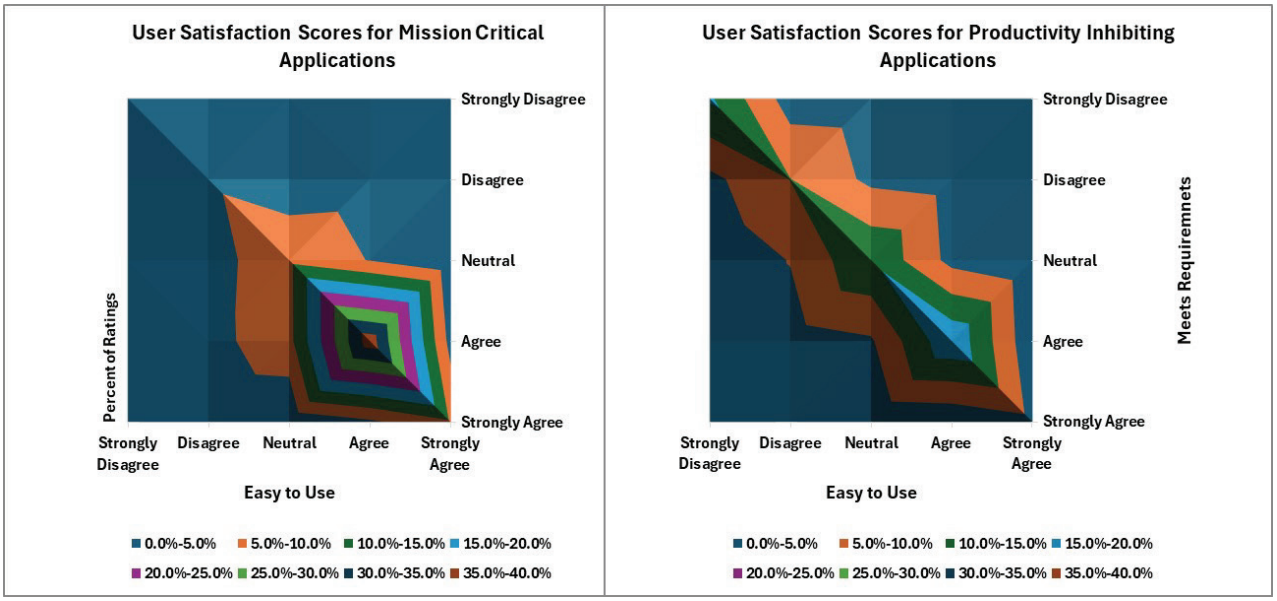
[13] Lewis, Utesch, and Maher, 2015.

[14] We did remove duplicate entries from a respondent.

- Respondents were more likely to rate browsers as easy to use than they were to rate them as meeting their needs; this effect is more pronounced when browsers were classified as a productivity-inhibiting application. This reflects users' complaints that browsers routinely block them from accessing sites containing information needed in the performance of their duties.

The average score for all mission-critical application ratings is 3.63 for meets requirements and 3.56 for easy to use. Applications that were rated as being productivity inhibiting have an average score of 2.74 and 2.75 on these two dimensions of user satisfaction, nearly a total point lower. Figure B.3 provides a side-by-side comparison of the distribution of these scores across the two dimensions for mission-critical applications (left panel) and productivity-inhibiting applications (right panel). While most ratings of the mission-critical applications fall within the range of neutral to strongly agree, the ratings for the productivity-inhibiting applications are more evenly distributed over the range from strongly disagree to strongly agree.

Figure B.3. Distribution of User Satisfaction Scores for All Application Ratings



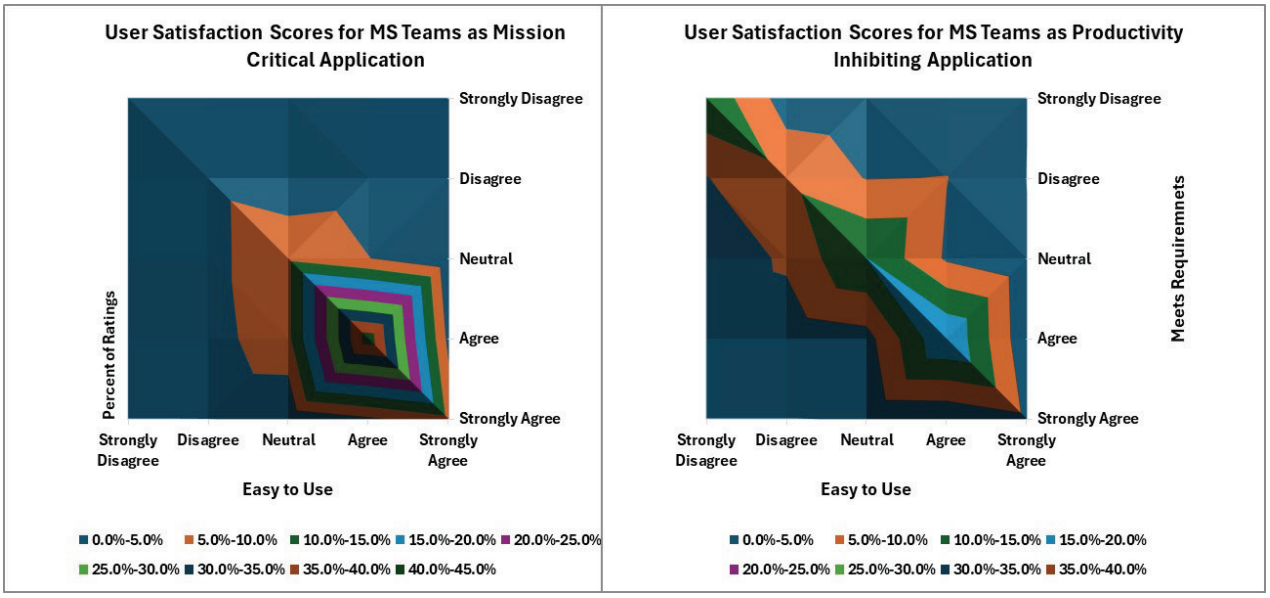
In many cases, the application being rated is the same. As we saw in Table B.4, how users rate Microsoft Teams depends on whether they see it as mission-critical application. This is shown in Figure B.4, which graphs Teams user satisfaction as a function of whether it was rated as a mission-critical application (left panel) versus a productivity-inhibiting application (right panel). For those who rated Teams as mission critical, they are relatively satisfied that it both meets their needs and that it is easy to use. For those who rated it as a productivity-inhibiting application, there is much less agreement that it meets their needs or that it is easy to use. In the responses to Question 19, which allows respondents to elaborate on their views, many of those who rated Teams poorly did so because their computers are configured to automatically start the application when the computer starts up. This automated start-up consumes significant computing resources and delays those users' ability to begin meaningful work. For those on intermittent networks or whose computing hardware is

inadequately sized for the software needed, the automated Teams start-up is a source of great frustration.

This frustration is the basis of our recommendation that local administrators review any applications that have an automatic start on computer turn-on and remove those that are not critical to the mission being performed. In particular, service personnel who share a general laptop among a dozen or more users (which both active duty and reservists tell us is common) should not have to wait for unneeded applications to start just so that they can check the current status of their orders or perform other administrative tasks. As one reservist explained:

> We are required to keep up with F2F, MyIMR, AROWS-R orders, MyFSS EPBs and other requirements, MyLearning CBTs and other annual requirements, ACES training stats, etc. And yet every month our unit does not have enough laptops/ desktops to perform these requirements during the drill weekend. Until this quarter, we had about 12 working computers for 120 personnel.[15]

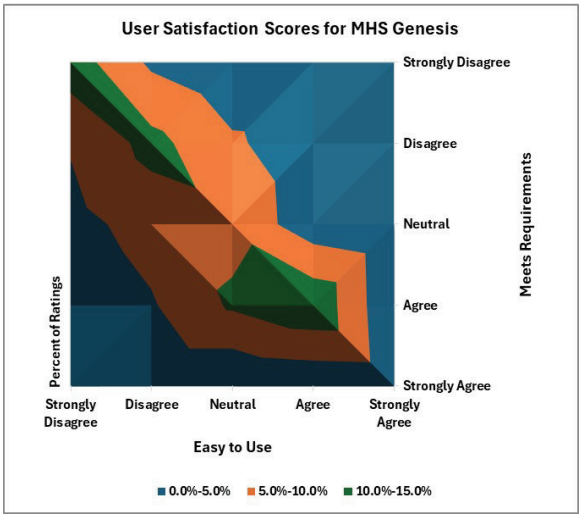
Figure B.4. Distribution of User Satisfaction Scores for Microsoft Teams



While the ratings for most mission-critical applications are concentrated in the neutral to strongly agree range, this is not true for MHS Genesis. The distribution of its ratings, shown in Figure B.5, looks very much like the distribution for a productivity-inhibiting application. In the open-ended comments regarding MHS Genesis, a recurrent theme was the inability to print the medical labels and wristbands essential to tracking specimens and patients in hospitals and labs. While printer problems may be a minor issue for many service members (and perhaps even for medical doctors), this is not

[15] Lack of definitions for these acronyms is in the original. Many IT and software applications share acronyms, and we are unable to identify the acronyms from the context provided. Our inability to identify the acronym does not change the point the respondent is making—they have many requirements that necessitate use of a computer but only limited opportunities to access a computer.

Figure B.5. Distribution of User Satisfaction Scores for MHS Genesis



true for the many the nurses and technicians who use MHS Genesis to perform their essential duties. As one explained:

> I currently have a IT ticket in because I cannot print wrist band labels from Genesis which we need for every recruit (about 800–1,000 a week) . . . it is beyond difficult to reschedule whole companies of recruits just frustrated with the slowness . . . we need help someone must know how to make everything work together[16]

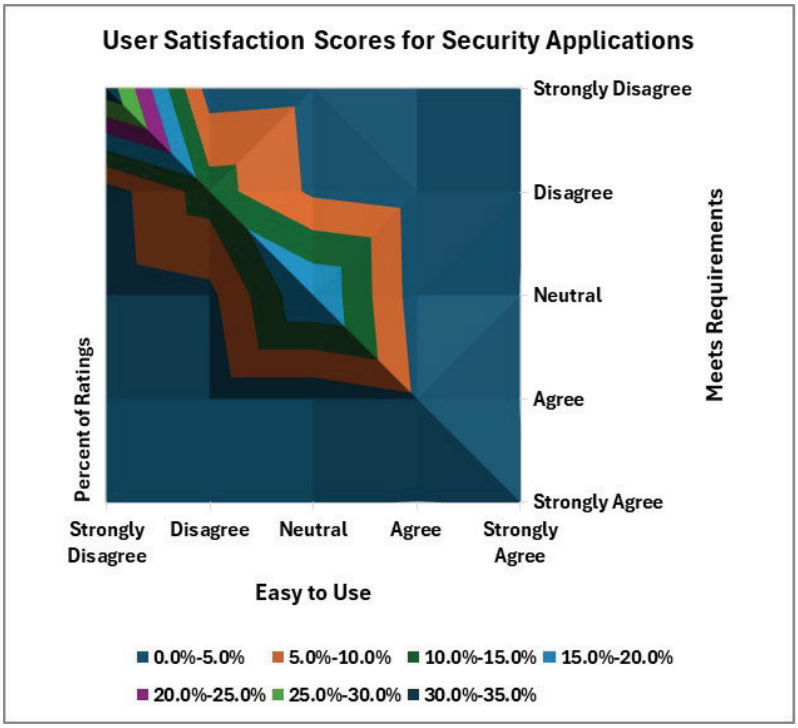
Another explained that in their operational setting, there were no directly connected printers and that constant network outages ("internet coverage in the hospital is atrocious") meant that just printing a document averaged 5 to 10 minutes. They concluded with the note that "As a medical provider, I spend at least 10–15% of my day addressing IT issues." As both respondents noted, deploying an application is not enough; the environment into which it is deployed must also be considered if the application is to be useful.

Among the productivity-inhibiting applications, respondents rated the various security applications quite poorly, especially those that consume significant computing resources, limit access to needed websites, or request credentials multiple times per a routine access. Less than 10 percent of those who rated these applications felt that they met their needs, and only 12 percent found them easy to use. As depicted in Figure B.6, 35 percent strongly disagreed with the statements that the applications met their needs or are easy to use. Too many times, we heard from users that a security application causes cascades of issues that lead to systems locking up or crashing. A service member described one of these cascades thusly:

> Constant [common access card (CAC)] pin prompts from every application hinders productivity. On average, I bet I enter my CAC 200 times daily. CAC pin window freezes application or forces you to halt one app or function to enter your pin in another area. Constant issue. For MS Teams and Alert, if the CAC pin or authentication window stays open too long, it locks up the virtual hypervisor running in background and there is no way to recover. A cold boot is required.

[16] Lack of punctuation is in the original.

Figure B.6. Distribution of User Satisfaction Scores for Security Applications



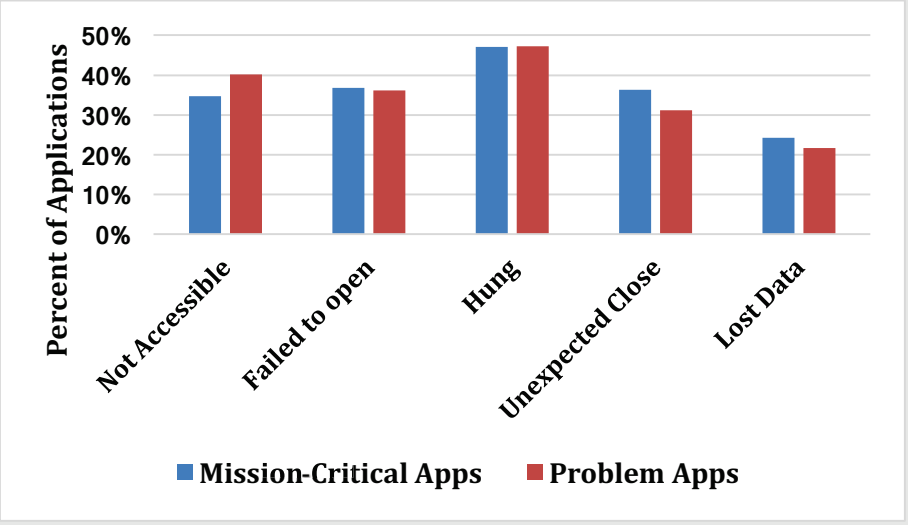
Q15. Issues Encountered in the Last Month

Interestingly, there is little difference in the type of issues encountered in the last month when comparing ratings for mission-critical applications with those that are productivity inhibiting. Figure B.7 provides a view of the percentage of applications reported as having encountered issues in the last month, by type of issue. While having an application hang is the most reported issue (~45 percent of applications), lack of accessibility, failures to open, and unexpected closures are nearly as prevalent at ~35 percent each. Issues that result in a loss of data are less prevalent—at around 20 percent—which is not surprising given that Microsoft Office applications are the most frequently rated and routinely provide auto-recovery files when those applications close unexpectedly.

Almost half of all applications have encountered multiple issues. On average in the last month, the applications reported here—both mission critical and productivity inhibiting—exhibited two of the five types of issues we inquired about, significantly underperforming industry norms for reliability of software applications.[17] Note that we did not inquire as to the frequency with which these issues occur, but from the open-ended comments; we know that some users encounter these issues multiple times a day. Although recovering from many of these issues can be relatively straightforward, the high

[17] The advent of application storefronts has provided a means to collect data regarding users' tolerance for poorly performing applications. Based on this data, several application developer websites state that the benchmark for a "good" or "3-star" application is fewer than 1 percent crashes per user, < 0.1 percent crashes per session, and <0.01 percent crashes per screen view. See, for example, Gary B, 2020.

Figure B.7. Issues Encountered in the Last Month



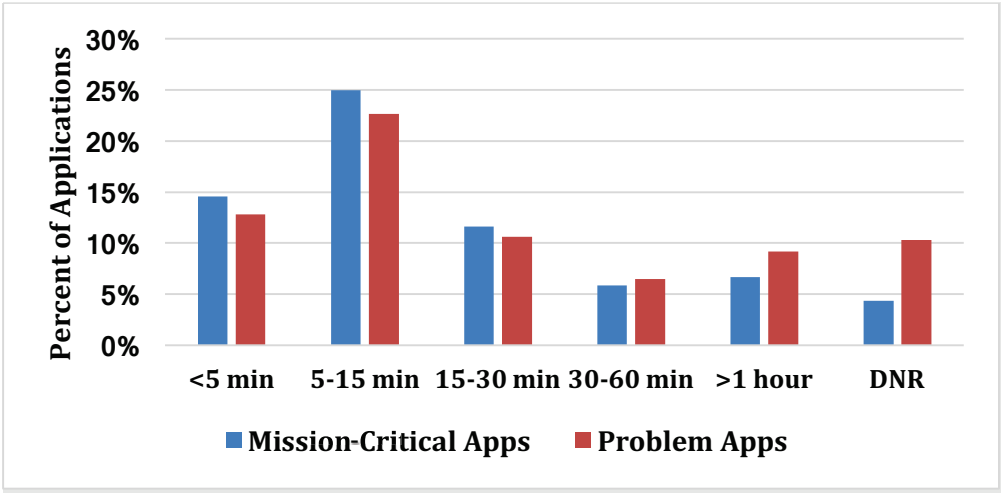
Note: Respondents chose all that apply; therefore, percentages do not sum to 100.

percentage of applications having rather severe issues is undoubtedly a factor driving perceptions of the inadequacy of DoD IT and application software.

Q16. Time to Recover from Most Recently Encountered Issue

It is in the time to recover from an issue that we see greater differentiation between the mission-critical and productivity-inhibiting applications. Figure B.8 shows that 26 percent of productivity-inhibiting applications and 24 percent of mission-critical applications take more than 15 minutes to

Figure B.8. Time to Recover from Issues



NOTE: DNR = did not recover.

recover from the most recent issue, a difference that is not statistically or practically significant. However, within those applications that take longer than 15 minutes to recover, a higher percentage of productivity-inhibiting applications take longer to recover, and 10 percent of productivity-inhibiting applications never recover at all (as compared with 4 percent for the mission-critical applications).

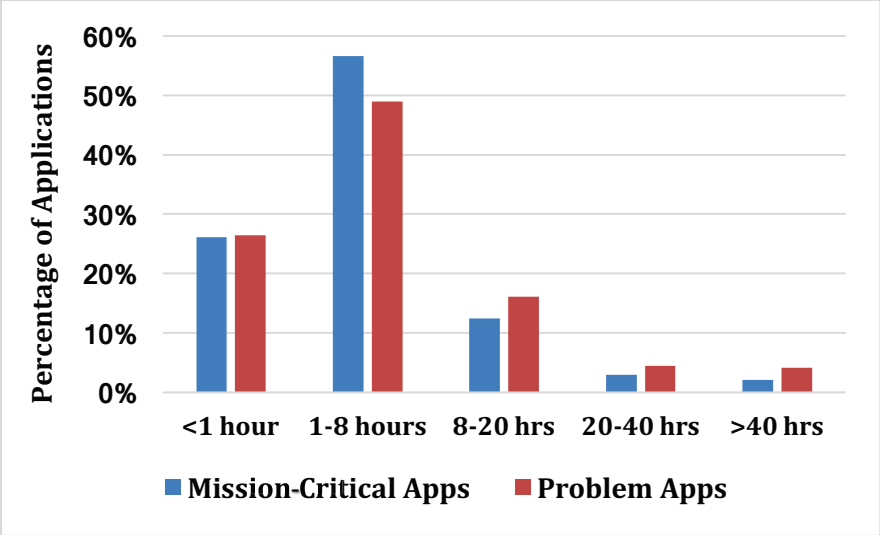
Q17. Usual Time to Recover from Issues

As with our other control questions regarding temporal biases, a majority of respondents (60 percent) indicated that the reported application time to recover after an issue was "about the same" as usual. For the remaining respondents, half stated the time was better, and half stated it was worse. There is no skew to the results. Therefore, we are comfortable saying that the results for application start-up time are relatively unbiased by temporal considerations.

Q18. Time to Regenerate Lost Work

The time needed to regenerate lost work also distinguishes mission-critical applications from productivity-inhibiting applications. Figure B.9 shows that while 83 percent of lost mission-critical work can be recovered in less than eight hours, only 75 percent of the work lost when productivity-inhibiting applications experience issues can be recovered in less than eight hours. In fact, when it takes more than 40 hours to regenerate work lost, it is twice as likely that the application experiencing the technical issues that caused the loss is rated as productivity inhibiting.

Figure B.9. Time to Regenerate Lost Work



Q19. Open-Ended Response

We received 1,875 valid responses to our question asking whether respondents wished to share anything else; this represents a response rate of 29 percent for Question 19.[18] Response rates were relatively even across the services, as shown in Table B.5, with the Navy the most likely to comment (34 percent). OCONUS users were slightly more likely to comment (32 percent) than CONUS users (28 percent).

Table B.5. Percentage of Respondents Who Commented

Service Affiliation	Surveys Started	Comments Received	Percentage Who Commented
Army	2,278	678	30%
Air Force	2,078	559	27%
Marines	388	95	24%
Navy	965	332	34%
Space	54	18	33%
Health	295	78	26%
Other	404	115	28%
Total	**6,471**	**1,875**	**29%**

We used a mixed method approach in our analysis of these comments. The comments were systematically coded at four levels of analysis:

- **primary:** perceived causes of underperforming technology
- **secondary:** perceived impacts on productivity or mission
- **emotional:** sentiments tied to both the perceived causes and perceived impacts
- **feelings of agency:** perceived ability to influence IT issues.

To remove the possibility of intercoder variation biasing our results, a single analyst coded for perceived causes of IT infrastructure and software issues (the primary coding), a second analyst coded for perceived impacts (the secondary code), and a third coded for emotions and feelings of agency.

Tabulating the raw number of comments for each of the codes provided a quantitative overview. The comprehensive analysis of primary and secondary codes highlights significant inefficiencies and dissatisfaction due to underperforming technology within DoD. We then correlated the emotional sentiments tied to these codes, categorized as positive, negative, or neutral, with subcategorizations to distinguish levels of positivity (satisfaction, optimism) or negativity (concern, confusion, annoyance, anger), which correlated to the perceived causes and perceived impacts. Finally, we reviewed the comments to determine whether the commenter perceived that they had any agency regarding how

[18] We discarded entries of "N/A," "None," or other expressions indicating a null response.

issues with IT are resolved; the codes were "has agency," "neutral," and "lacks agency." A combination of strong negative sentiments and lack of agency regarding the source of those sentiments has been correlated with a higher likelihood that a worker will leave a position.[19]

Primary Coding Results: Perceived Causes

We coded each comment to reflect our interpretation of the cause(s) a respondent appears to blame (or in rare cases, praise) for technology-related issues. Comments often address multiple perceived causes. We developed the primary codes on the basis of prior research as well as inductively during the data analysis. See Table B.6. for a brief description of the codes used.

Table B.6. Codes for Perceived Causes of Information Technology–Related Issues

Primary Code	Description
Hardware	Physical components of IT systems, including computers and peripherals, crucial for executing software and network operations
Network access	The ability to connect and interact with data networks, including via non-classified internet protocol routers, secret-level internet protocol router, and the Joint World-Wide Intelligence Communication System
Software	Applications and programs used by DoD personnel to perform tasks, including operating systems and tools such as Adobe Acrobat and the Microsoft Office Suite
Imbalance	Disparities in compatibility between software applications and hardware components, which can lead to inefficiencies and hinder overall system functionality
Policy	Guidelines and regulations governing the use, security, and maintenance of IT systems within the DoD
Training	The provision of knowledge and skills to DoD personnel to effectively utilize IT systems and software
Support	Assistance provided to troubleshoot, maintain, and optimize IT systems and applications for DoD personnel
Updates/upgrades	Enhancements or patches applied to software and hardware to improve performance, security, and functionality
Licensing	Authorization and management of software usage rights and subscriptions necessary to access and use specific applications

NOTE: DoD = Department of Defense; IT = information technology.

These primary codes encompass a range of perceived causes, such as hardware limitations, network access issues, software inadequacies, imbalances between hardware and software, and problematic policy implementations. In additional, the codes include training deficiencies, support challenges, issues with updates and upgrades, and licensing problems. Each primary code helps to identify specific areas where improvements may be needed to enhance IT functionality and overall productivity.

[19] Thomas A. Wright and Russell Cropanzano, "Emotional Exhaustion as a Predictor of Job Performance and Voluntary Turnover," *Journal of Applied Psychology*, Vol. 83, No. 3, 1998.

Table B.7. details the distribution of these primary codes across DoD. Differences between services are not statistically significant, the one exception being the Navy's complaints regarding network access. At the time of the survey, the recent loss of the Navy's primary VPN had left many Navy users without remote access.

Table B.7. Perceived Causes of Information Technology–Related Issues

Primary Code	Army	Air Force	Marine Corps	Navy	Space Force	Health	Other	*TOTAL*
Software	47% *(317)*	48% *(262)*	42% *(39)*	44% *(144)*	— —	53% *(41)*	46% *(53)*	*35%* *(864)*
Hardware	37% *(249)*	46% *(255)*	58% *(53)*	44% *(145)*	71% *(12)*	42% *(33)*	38% *(43)*	*32%* *(790)*
Policy	20% *(135)*	15% *(82)*	— —	17% *(57)*	— —	— —	18% *(21)*	*13%* *(315)*
Network access	8% *(54)*	5% *(26)*	— —	17% *(57)*	— —	— —	— —	*6%* *(148)*
Support	7% *(44)*	6% *(33)*	— —	7% *(23)*	— —	— —	9% *(10)*	*5%* *(123)*
Updates/Upgrades	7% *(47)*	3% *(17)*	— —	4% *(12)*	— —	— —	— —	*4%* *(88)*
Imbalance	4% *(24)*	4% *(22)*	— —	— —	— —	— —	— —	*2%* *(61)*
Training	2% *(16)*	3% *(17)*	— —	— —	— —	— —	— —	*2%* *(45)*
Licensing	1% *(10)*	2% *(11)*	— —	4% *(12)*	— —	— —	— —	*1%* *(37)*

NOTE: Codings with fewer than ten responses are omitted from the table but are included in the totals; therefore, not all columns will sum to 100 percent. When adjusted for sample size, differences between services are not statistically significant, with one exception—the Navy's complaints regarding network access. Since a comment can contain more than one perceived cause, percentages are of all perceived causes rather than the percentage of all comments.

Software inadequacies were the most cited perceived cause of technical issues, at 38 percent of all primary codings. Common issues included non-intuitive interfaces, frequent application crashes, and the inability to access critical software tools. Applications such as Defense Ready and Teammate+ were specifically mentioned as problematic in that they often require multiple restarts and cause significant work disruptions. DoD personnel reported that these software-related problems were substantial impediments to productivity. One wrote, "Core software like Outlook and Teams seems to have good investment. However, ancillary systems . . . suffer from major lack of investment and are overall a broken process."

Hardware limitations were cited almost as frequently as software limitations, at 32 percent. Outdated equipment, insufficient random-access memory, and slow performance were recurrent themes. Some respondents reported using personal laptops with advanced specifications to mitigate these limitations, highlighting the stark inadequacy of the provided hardware. These issues were

critical barriers to efficient work, resulting in substantial delays and reduced productivity. One commenter wrote:

> Countless manhours are lost daily across the DoD due to inadequate software, but mostly it is inadequate hardware. . . . With current security protocols 32 GB [gigabytes] of RAM [random-access memory] is the absolute minimum needed to do basic things. I often bring in a personal laptop with a CAC [common access card] reader that is hot spotted to my phone to perform [DoD] work as it is lightyears faster than any computer or network [DoD] ever provided to me.[20]

Another commenter summed it up succinctly: "The computers are simply slow. They are slow to boot up, slow to open files and slow to access networks."

After software and hardware, the next most cited cause of technical issues is policy, at 13 percent. Policy-related challenges include restrictive cybersecurity measures and inconsistent software standards. Some respondents noted that abrupt policy changes restricted access to software without providing viable alternatives, thereby necessitating cumbersome workarounds. One respondent told us:

> Security is of course a top priority, but [we] are not able to fully meet mission requirements because security is so stringent. I understand the need for secure systems, but unfortunately our existing systems are so secure that I don't use them, instead electing to use my personal devices where I can actually get work done. Some kind of compromise is needed.

Network access and IT support are the next most frequently cited causes of technical issues, at 6 and 5 percent, respectively. Network access issues include frequent VPN dropouts and intermittent network performance. The persistent and widespread nature of these connectivity problems cause significant downtime and hamper the ability to complete essential tasks. A typical comment reads: "Connectivity is the biggest issue. VPN connection constantly dropping. I had to reload this survey 5 times to complete it." VPNs were the sixth most rated mission-critical application in our survey, indicating their importance to the mission. Given that two services lost access to their primary remote access solution recently, we recommend that DoD improve the resiliency of these offerings.[21] Many respondents indicated experiencing long resolution times when seeking support for technical problems. Delays in receiving technical support and frequent miscommunication add to the frustration, especially for those who struggled to receive notice of the status of their issues without having network access or a working computer. One commenter summed up the intersection of these two causes: "Connectivity is slow, tech support is lacking, can't put in tickets because the system is down." Another noted:

> I also requested a phone call for any update on my ticket because I could not access my computer email and no contact was made. I had to call every day for an update and they would say, I sent you an email. I would have to remind them that I could not log on to the computer. . . . [There needs to be] an easier solution for us remote workers with system/IT issues.

---

[20] In extracting quotations from the survey comments, we have edited out information that reveals a commenter's service affiliation, location, rank, or position.

[21] In addition to the Navy's loss of their primary VPN, the Air Force had also run into issues with their remote access offering at the time the survey was conducted.

Unfortunately, there were other similar complaints, and one user told us of having to drive over 600 miles to the nearest support center multiple times to have their issue resolved.

The next most cited cause for technical issues is the software update process. Multiple users asked why updates that force them to reboot their computers happen in the middle of their shift. As more than one user lamented, "Why can't these updates occur on the weekend?"[22] Another source of complaint is upgrades that render systems unreliable. Respondents told us that poorly thought-out upgrades often lead to system instability, necessitating lengthy tech support interventions and causing significant productivity losses. As we will note later when discussing correlations between emotions and causes, a poorly perceived software update is highly correlated to feelings of anger (not just frustration).

We also coded to better understand how well users understand that it is the imbalance of inadequate hardware resources relative to the demands of the software that drives technical issues. In our review of the productivity-inhibiting applications, we had noted that 40 people had cited "bloatware" as an issue and wondered if the comments could give us more insight into this issue. Just over 2 percent of the primary coding highlights issues in which software demands outstripped the capabilities of outdated hardware, leading to slow performance and frequent system crashes.

Training issues were also cited in 2 percent of the causal codings. Some personnel reported insufficient guidance on navigating complex software applications, leading to inefficiencies and frustration. They suggested that enhanced training programs would be a necessary step to equip DoD personnel with the skills needed to maximize the potential of the provided IT resources.

Comments also included licensing issues, such as the lack of access to necessary software due to unapproved or unaffordable licenses (although these comments represent just 1 percent of the primary coding). Personnel often cope with free versions of essential applications as opposed to more costly features-rich versions, and these limitations hinder their ability to complete tasks efficiently. They regard addressing licensing problems to ensure access to required software as vital for improving productivity.

Secondary Coding Results: Perceived Impacts

Secondary codes captured the specific impacts of these technology issues on productivity. Table B.8 shows the codes we identified and used.

Table B.8. Codes for Perceived Impacts of Information Technology–Related Issues

Secondary Code	Description
Does not meet need	IT systems and software fail to fulfill the requirements of DoD personnel, hindering their ability to perform tasks effectively and execute their mission.
Hard to use	IT applications and systems have cumbersome, non-intuitive interfaces that make them difficult to navigate and operate.
Time waster	Significant delays in productivity are caused by slow system performance, extended login times, frequent software crashes or updates, and the like.

[22] However, as a reservist pointed out to us, performing all updates on the weekends would severely affect their training time. There is no universal "best time" to make software updates, and we encourage DoD to use tools that allow users to self-schedule updates within the next 24 or 48 hours. While urgent cybersecurity patches may need to happen on a more immediate basis, it seems that many updates are causing unnecessary disruption.

Secondary Code	Description
Implementation of security policy	Challenges related to IT policies and security measures lead to inconsistent software standards and disrupted workflows.
Lack of network access	Inadequate network connectivity and unreliable VPN connections prevent personnel from accessing necessary applications and data efficiently.
Lack of tools	Absence of essential software and licenses hampers the ability to complete tasks.
Lack of interoperability	Incompatibility between different IT systems and software causes frequent communication and functionality issues.
Bad update/upgrade	Poorly executed system upgrades disrupt productivity and/or lead to decreased system performance.

NOTE: DoD = Department of Defense; IT = information technology; VPN =virtual private network.

We used the secondary codes to categorize the perceived impacts of underperforming technology on the productivity of DoD personnel. These codes capture a comprehensive range of issues, including hardware and software failures, network access challenges, and policy-related inefficiencies. By analyzing responses through these secondary codes, we were able to assess perceived impacts of underperforming IT on daily operations for DoD personnel. Table B.9. below details the distribution of these secondary codes across DoD, highlighting the frequency and proportion of each issue as related to the total number of responses.

Table B.9. Perceived Impact of Information Technology–Related Issues

Secondary Code	Army	Air Force	Marine Corps	Navy	Health	Other	*TOTAL*
Does not meet need	10% *(104)*	9% *(79)*	14% *(20)*	12% *(61)*	12% *(16)*	15% *(25)*	*11% (309)*
Hard to use	9% *(96)*	6% *(52)*	10% *(14)*	11% *(55)*	10% *(13)*	9% *(16)*	*9% (250)*
Time waster	24% *(253)*	28% *(239)*	23% *(32)*	21% *(112)*	30% *(40)*	22% *(37)*	*25% (718)*
Implementation of security policy	19% *(195)*	17% *(146)*	16% *(22)*	14% *(73)*	13% *(17)*	16% *(27)*	*17% (488)*
Lack of network access	23% *(242)*	23% *(198)*	24% *(33)*	24% *(127)*	23% *(30)*	18% *(31)*	*23% (666)*
Lack of tools	9% *(93)*	9% *(79)*	11% *(15)*	12% *(63)*	—	14% *(24)*	*10% (283)*
Lack of interoperability	3% *(27)*	1% *(10)*	—	2% *(10)*	—	—	*2% (57)*
Bad update/upgrade	4% *(41)*	5% *(43)*	—	4% *(20)*	—	6% *(10)*	*4% (122)*

NOTE: Space Force comments are not broken out separately due to the small sample size but are included in the totals. Codings with less than ten responses are omitted from the table; therefore, not all columns will sum to 100 percent.

Does Not Meet Need

Eleven percent of respondents state quite plainly that the existing IT infrastructure does not adequately meet their needs and negatively affects their ability to execute their mission. For instance, when describing the impact of a licensing issue, a commenter noted that "the software provided does not meet the overall need of [my] position." Another respondent echoed this sentiment, while noting challenges encountered by reserve members and remote workers: "As a Reserve officer, the lack of a meaningful way to remotely connect to [the network] impacts my ability to complete tasks."

Hard to Use

The usability of key DoD software was a complaint of 9 percent of the commenters. Respondents describe software as cumbersome and non-intuitive, requiring multiple restarts for basic functionality and significantly reducing work efficiency. As one commenter put it, "Everything requires multiple hoops to jump through just to get the most basic tasks accomplished." Difficulty using IT applications was cited more frequently as impacts on productivity than they were as the perceived primary cause of productivity challenges. As one respondent stated, "New software must be easy to use 'out of the box.'" Another noted that even training websites are not user-friendly, saying, "The majority of training websites are not easy to use, which makes it hard to complete them in a timely fashion."

Time Waster

One quarter of comments cite delays in performing tasks as an impact of their IT-related issues. They describe extensive time lost due to slow computer boot times, sluggish application performance, and frequent software updates. One commenter noted that "it is normal for simple tasks to take between two to four hours, if the system is up and running." Another frequently mentioned issue and source of frustration was time spent just logging on and dealing with forced updates. So were persistent common access card PIN entry requirements, which respondents saw as slowing workflow considerably and exacerbating inefficiencies.

Implementation of Cybersecurity Policy

Seventeen percent of comments concern the choices DoD organizations have made when implementing cybersecurity policy and related measures. For example, a user explained how cybersecurity policies that restrict the use of older Microsoft Office versions prevented them from archiving and accessing information regarding legacy systems without providing suitable alternatives. The net result was that the user was unable to efficiently (or perhaps even inefficiently) meet their mission requirements. One individual summarized the feelings of many of the commenters in saying that "installed computer security systems do a poor job of balancing security and access."

Another implementation-related concern highlighted during the secondary coding is the impact of duplicative cybersecurity scans that consume many of the computing resources. Users perceive these as impediments to productivity and ask for a better balance in their implementation. As one commenter explained,

> Duplication of security tools performing scans at the same time consumes valuable resources that are required to perform a job function. DoD continues to implement more cyber tools without determining if there is an existing tool performing the same function.

Inconsistent implementation of policies regarding access controls for information and data is another issue, as is the associated impairment of mission tasks. According to one commenter,

> The primary issue I have . . . is the inconsistent and incomprehensible way that information access is managed and controlled. There is generally no available guidance . . . making it subject to individual willingness to allow access.

Another echoed this sentiment, stating that "as someone that needs access to external content [to perform job duties], I need web controls that are consistently applied."

Variations on this theme came from lawyers who could not access case files, medical professionals who could not access medical journals, instructors who could not access their curricula, Army officers who could not access Army websites, and the many commenters who told us they could not access our survey from their assigned end-user devices.

Lack of Access

Network connectivity was a major concern, with 23 percent of comments highlighting this aspect. Unreliable VPN connections, intermittent local area network performance, and overall poor network quality all contribute. As DoD moves to allowing more remote work, this has particularly large impacts. As one leader noted, "VPN issues have slowed the productivity of remote workers (60% of my team)."

However, in-office personnel, regardless of whether they are CONUS or OCONUS, may not be faring much better in terms of the negative impacts of network connectivity issues on productivity. While a stateside respondent stated that "service is actually worse when working in the office, while hard-wired to the network," an overseas respondent said that "the network goes down a lot. There are times when the network is down for days."

In addition, many respondents reportedly experienced significant downtime before connectivity issues were resolved. For example, one commenter stated that "as a whole, our network has outages where no IT-based work can occur . . . bringing the work environment to a near complete halt."

Lack of Tools

Respondents repeatedly mentioned a lack of essential tools, such as full Adobe Acrobat Pro software licenses for digitally signing PDFs. Personnel often had to cope with read-only versions or lacked access to critical software, which severely affected their ability to complete tasks efficiently. One commenter stated that

> Adobe is a particular problem. The license is always out of date, and you can't fill anything out when it's expired. This can basically stop all work on a project when this happens.

The bottom line, as one said, is that "lack of software is limiting my organization's ability to do our jobs."

Lack of Interoperability

Commenters reported interoperability issues between hardware and software as substantial impediments. One respondent stated that "software and IT should be interoperable. Systems don't talk to each other, and it slows down the work and how efficiently we can get tasks done." Another

frequently mentioned issue was the lack of interoperability between networks, such as between combatant command networks and military department networks, which impedes communication and has a direct impact on operations. As one said, "The primary issue from my experience is the lack of interoperability between networks."

Bad Update/Upgrade

Poorly executed upgrades were a significant source of discontent. Forced updates during work hours and updates that rendered existing systems unreliable were common issues. While civilians and active-duty military asked for updates to occur on weekends or other off-peak times because, as one commenter put it, "Automatic updates during working hours negatively affect the already slow network and hinders productivity," reservists rightly pointed out that those times would interrupt their use:

> In the Reserve, having [IT] updates over the weekend means the entire system is down for our entire work weekend. That is the same as if a system was down for an entire month on active duty.

Respondents noted that it is not simply that the software updates take time, but that they also often led to system inoperability, which in turn necessitates lengthy tech support interventions. As one respondent lamented, "My computer updated, even though I didn't want it to, and now my VPN doesn't work." Many users expressed frustration and even anger when their computers updated without notice, causing cascading impacts that too often led to hours of outage.

Emotional Coding Results

The emotional coding identified (a) the general emotion the comment conveyed and (b) the sense of personal agency the comment conveyed.

Coding by Emotion

Some comments conveyed positive emotions (satisfaction or optimism), some were neutral, and others conveyed negative emotions (concern, confusion, annoyance, frustration, or anger). These emotions are described in Table B.10.

Table B.10. Emotion Coding Analysis

Emotion	Description
Satisfaction	Pleased with the state of Department of Defense IT (e.g., makes positive statements about IT support received, makes positive statements about computing equipment)
Optimism	Expresses confidence in the direction of Department of Defense IT; may articulate minor issues, but overall feels things are going in the right direction
Neutral	Conveys neither clear positive nor negative emotions
Concern	Conveys worries (e.g., about the functionality of software, of impacts of IT on the defense enterprise)
Confusion	Conveys a lack of understanding (e.g., about IT in general, about how to use specific software applications)

Emotion	Description
Annoyance	Expresses displeasure with DoD IT but articulates problems in a matter-of-fact way
Frustration	Expresses strong displeasure with DoD IT, using emotional words (e.g., "painful," "ludicrous," "ridiculous," "awful," "crippling," "horrible") to describe problems and associated impacts
Anger	Similar to frustration, but even more strongly negative and emotional; often included multiple exclamation points, all caps, and/or swearing for emphasis

NOTE: DoD = Department of Defense; IT = information technology.

Table B.11 shows the percentage and comments that exhibited each emotion by service and for the total force. Overall, frustration was the most prevalent at (50 percent) emotion among those who answered Question 19, followed by annoyance (at 24 percent).

Table B.11. Emotion Coding Results Across Department of Defense Personnel

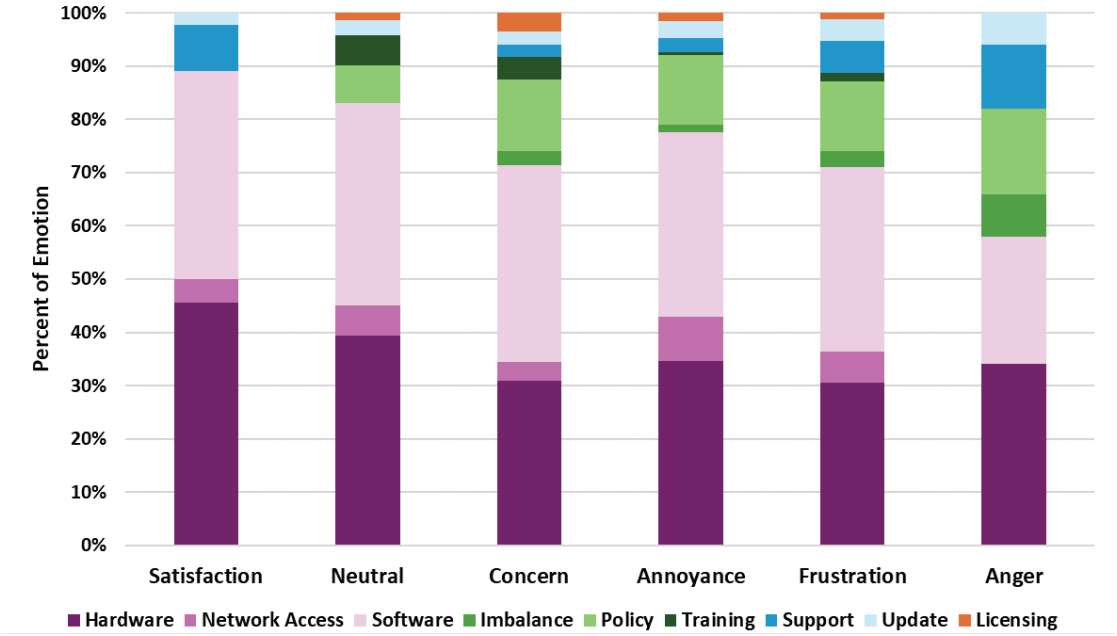
Emotion	Army	Air Force	Marine Corps	Navy	Health	Other	*TOTAL*
Satisfaction	3% (22)	2% (11)	—	—	—	—	3% (54)
Neutral	4% (28)	5% (27)	—	5% (17)	—	—	5% (86)
Concern	11% (77)	11% (64)	15% (14)	7% (23)	—	—	10% (196)
Confusion	2% (13)	—	—	—	—	—	1% (21)
Annoyance	26% (176)	23% (128)	19% (18)	21% (70)	23% (18)	26% (30)	24% (446)
Frustration	51% (348)	56% (314)	55% (52)	61% (204)	56% (44)	51% (59)	55% (1,030)
Anger	2% (13)	—	—	—	—	—	2% (34)

NOTE: Codings with fewer than ten responses are omitted from the table; therefore, not all columns will sum to 100 percent. Space Force comments are not broken out separately due to the small sample size but are included in the totals. Optimism was found in only five comments.

Figure B.10 shows the emotions by the perceived causes described in the "Primary Coding Results" section, above. While software and hardware are the leading causes associated with all emotions, policy issues are associated primarily with negative emotions, and the perceived imbalance of hardware and software is associated primarily with anger.

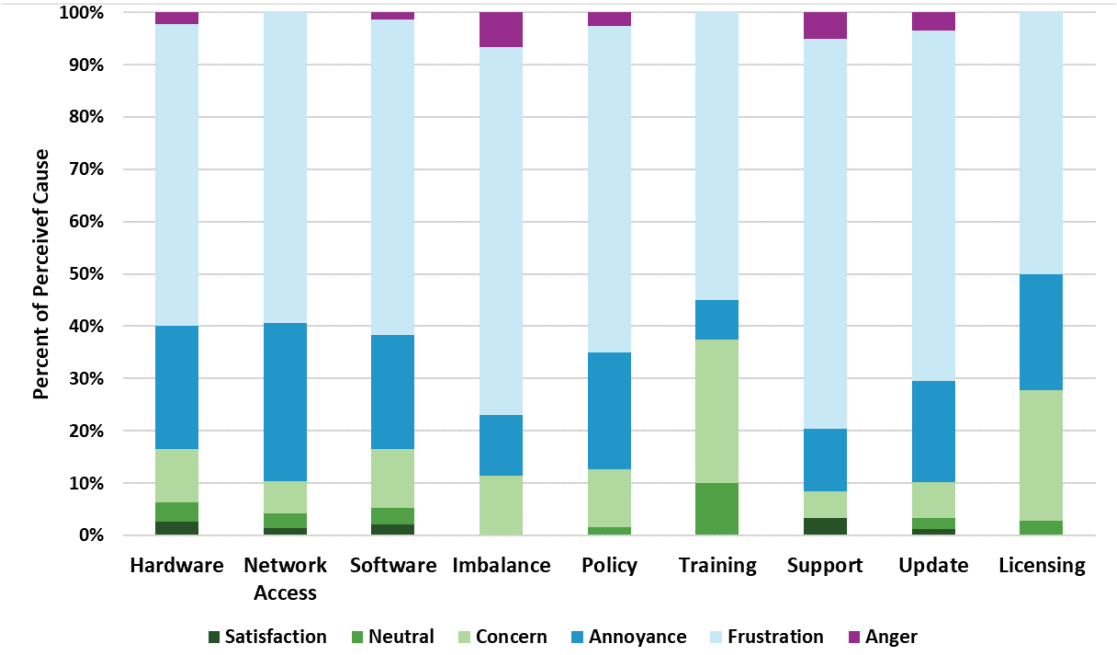
Figure B.11 shows the perceived causes by emotions (a different view of the same data). Note the disproportionate percentage of frustration and anger associated with the imbalance of hardware resources and software consumption of those resources. The emotions associated with support have a similar disproportionate percentage of frustration and anger.

Figure B.10. Emotion by Perceived Cause



NOTE: Optimism and confusion are omitted because they were identified only in a small number of comments.

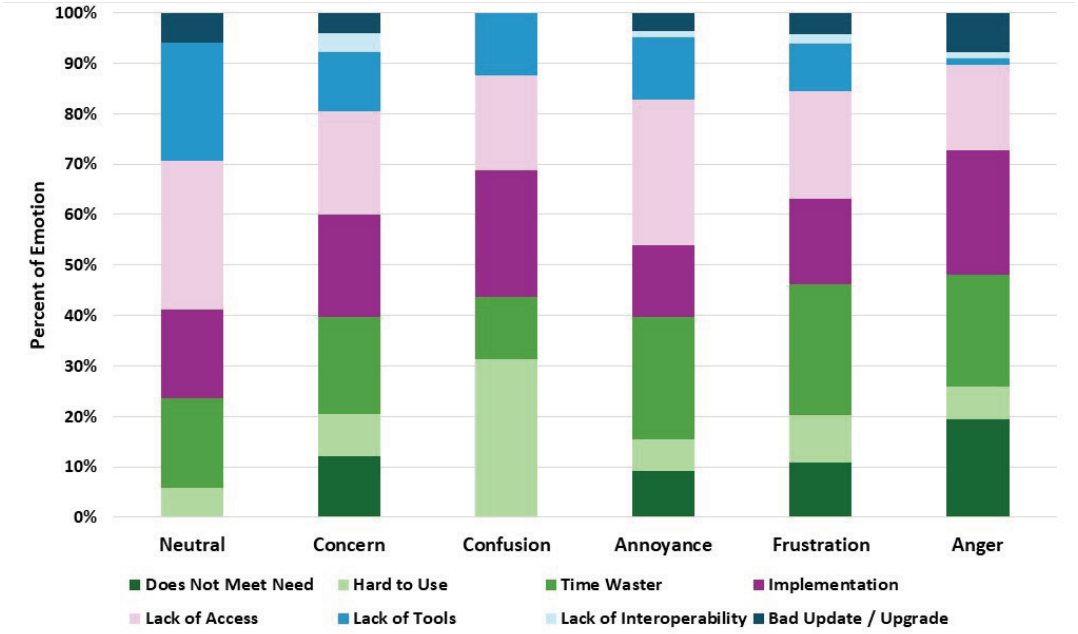
Figure B.11. Perceived Cause by Emotions



NOTE: Optimism and confusion are omitted because they were identified only in a small number of comments.

Figure B.12 shows the emotions connected with the perceived impacts described in the "Secondary Coding Results" section, above. Respondents particularly associated frustration with wasted time and lack of network access.

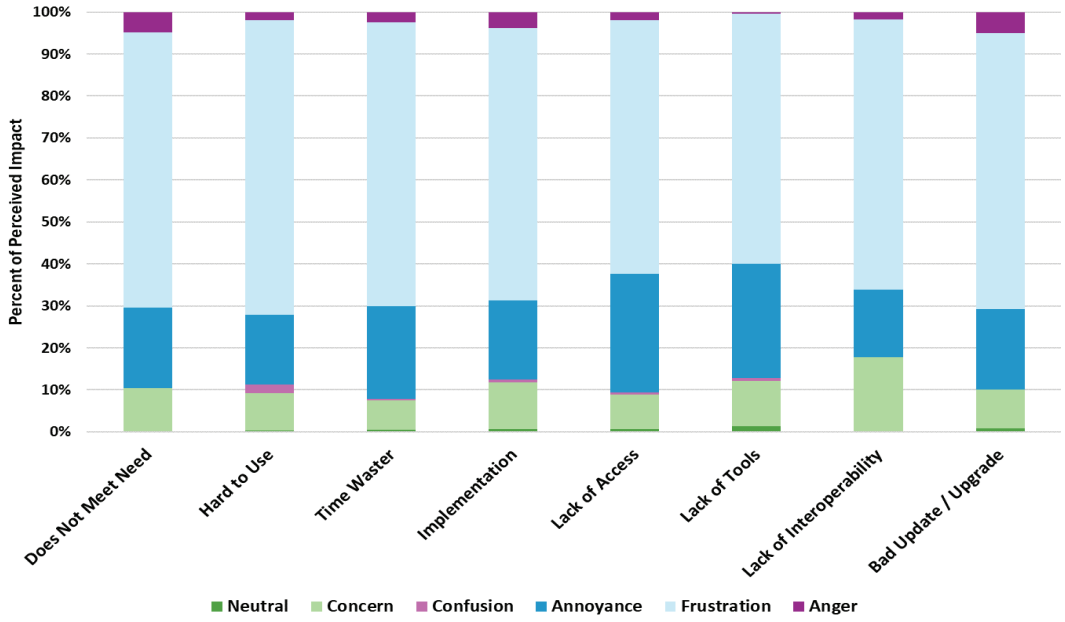
Figure B.12. Emotion by Perceived Impact



NOTE: Optimism and satisfaction are omitted because they appeared in this subset of data fewer than ten times.

Figure B.13 shows the perceived impacts by emotions (a different view of the same data).

Figure B.13. Perceived Impact by Emotion



NOTE: Optimism and satisfaction are omitted because they appeared in this subset of data fewer than ten times.

Coding for Sense of Personal Agency

We also sought to identify the extent to which those who responded to Question 19 exhibited signs of burnout, and particularly, emotional exhaustion, which could put them at risk of leaving their jobs. Multiple studies have defined three dimensions of burnout: emotional exhaustion, which "describes feelings of being emotionally overextended and exhausted by one's work";[23] depersonalization, a negative, cynical perception of others (e.g., clients); and diminished personal accomplishment, including "the tendency to evaluate oneself negatively."[24] Because researchers have found that emotional exhaustion in particular is a key driver of burnout,[25] that is what we focused on. To identify potential emotional exhaustion, we examined the relationship between emotion and sense of personal agency, where *agency* in this context refers to a subjective feeling "of control over actions and their consequences."[26] While strong agency is associated with positive feelings, such as confidence, satisfaction, and happiness, weak agency is associated with negative feelings, such as insecurity, fear, and exhaustion.[27]

We assigned comments one of the following three codes:

- Able to have impact: Just 2 percent of commenters expressed confidence that their actions influence outcomes (e.g., their work makes an impact, they can get IT support to resolve their issues when needed, and so on).
- Neutral: Thirty-two percent of commenters did not clearly indicate an ability or inability to influence outcomes.
- Unable to have impact: The majority of commenters (66 percent) expressed a lack of ability to influence outcomes (e.g., conveys feeling that no one can or will help, hopelessness, and so on).

Those who expressed both an inability to have an impact and the more negative emotions (frustration and anger) are likely at risk of quitting their jobs. We show the correlation of emotion to agency in Figure B.14.

While only 2 percent of commenters indicated anger, it is important to note that those who did also indicated that they feel an inability to affect outcomes. The many commenters (55 percent) who expressed frustration also feel that they are unable to change the way things are. This is a sizable percentage of the respondents who left comments. We know that the ~30 percent of respondents who commented have self-selected in by electing to tell us about their experience, perhaps in a last-ditch effort to have an impact. Furthermore, we need to account for the self-selection bias of those who responded to the survey at all. Therefore, these results do not mean that ~15 percent (i.e., half of the 30 percent of respondents who commented) of DoD personnel are at risk of departure. Accounting

[23] Wright and Cropanzano, 1998.

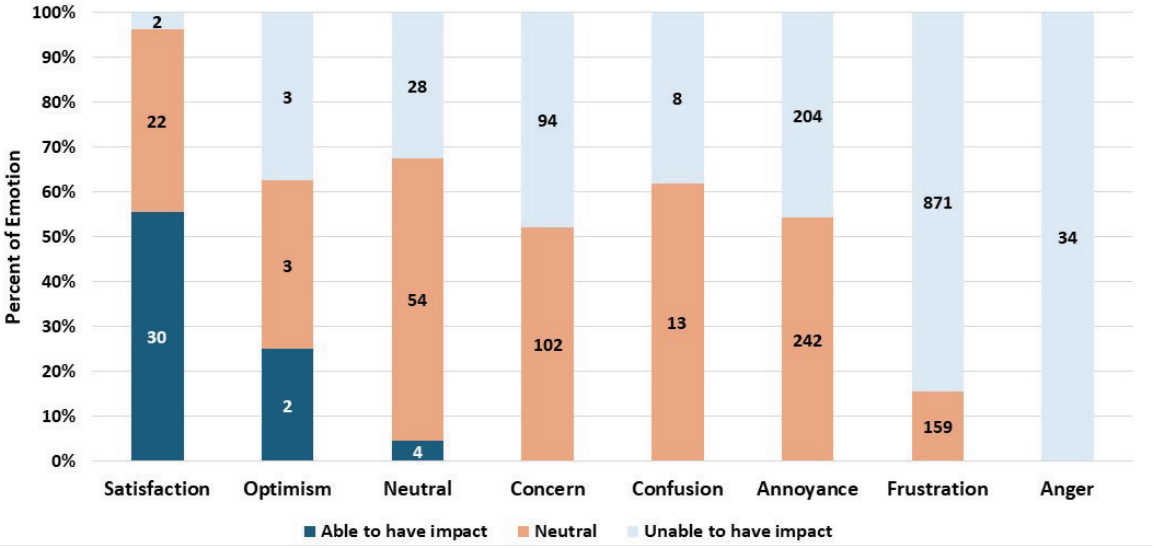
[24] Christina Maslach and Susan E. Jackson, "The Measurement of Experienced Burnout," *Journal of Occupational Behaviour*, Vol. 2, 1981.

[25] Wright and Cropanzano, 1998.

[26] Moore, 2016.

[27] Jani Ursin, Katja Vähäsantanen, Lynn McAlpine and Päivi Hökkä, "Emotionally Loaded Identity and Agency in Finnish Academic Work," *Journal of Further and Higher Education*, Vol. 44, No. 3, 2020.

Figure B.14. Emotion by Feelings of Personal Agency



for self-selection bias is not a science but, given the technical barriers we know respondents had to overcome to respond to our survey, we assume that only the most motivated did so and discount this measured result by a factor of three when applying to the total workforce. *Therefore, we conservatively estimate that 5 percent of the DoD workforce may be strongly motivated to depart from service due to poorly performing IT and software.*

Appendix C

Operating the Recommended Framework

The recommended Phase 1 framework exposes how different systems and information environments affect user experience and efficacy in accomplishing their mission using IT and software. The general idea incorporated in this framework is to systematically collect, report, and compare metrics regarding the usability and usefulness of systems along with the time consumed by client, network, and server processing in dynamically selected user groups. Because the user groups operate their systems in different information environments, differences in user-reported usability and usefulness and measured time consumed by client, network, and server processing expose the source, degree, and mission impact of friction in user interactions with a system. By measuring user experience at the end-user device and with the users themselves, this framework supersedes an individual system-by-system view and reports on the overall usability and usefulness of the entire chain of systems and devices that enable users to execute their mission using IT and software.

This appendix details the metrics to be collected in a Phase 1 implementation of the recommended framework and offers use cases to demonstrate how those metrics can inform decisionmakers throughout DoD. It introduces the data first and then describes some of the ways in which that data can be analyzed to identify friction in user interactions with their IT and software systems. Though the data used in this framework include performance metrics and budgetary information, it is not the intent of this framework to troubleshoot technical issues or provide cost-benefit analyses. Instead, the framework signals the source, degree, and mission impact of technical problems and helps identify outliers with respect to cost versus performance.

The Data Layer

To enable assessment and comparison of different programs, systems, and IT environments across DoD, all organizations will need to collect and expose a minimal set of data in a consistent manner. The framework presents a two-dimensional graph in which data from selected user groups are plotted on a value axis and a monetary axis. The heterogeneity of systems and IT environments within DoD permits comparisons of user experiences in a variety of information environments.

The Value Axis

In Phase 1, the value axis shows user satisfaction and/or technical performance ratings for dynamically selected systems of interest, collected by each organization and exposed to DoD.

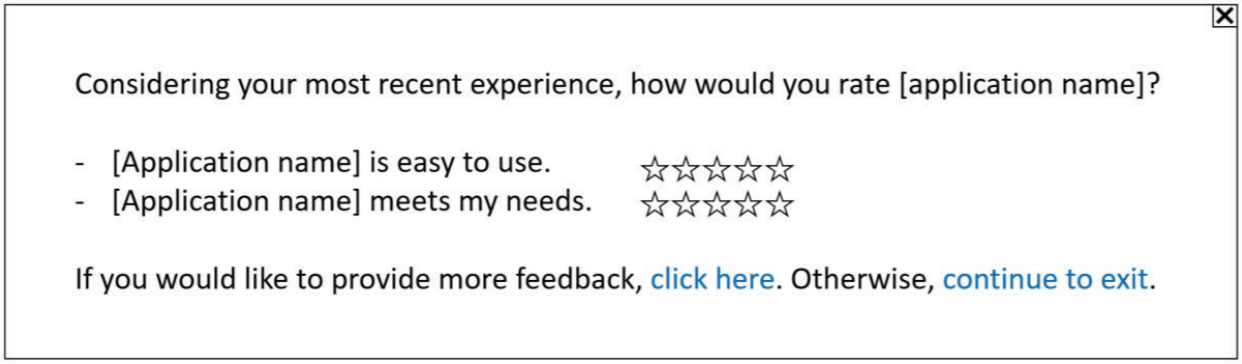
User Report

A large proportion of the department's workforce is technologically enabled to contribute to operational and institutional missions through IT and software systems provided by DoD. Consequently, the users' reports of inaccessibility or insufficiency of their IT and software tools provide essential insights into their ability to accomplish the mission. User report data in the framework consists of accumulated responses to two statements, collected for each user-facing application. These statements come from UMUX-Lite and refer to a user-facing system of interest:

- "[Application name] meets my requirements."
- "[Application name] is easy to use."

Each statement is presented to the user for a response on a scale from 1 (strongly disagree) to 5 (strongly agree), represented intuitively by one to five stars.

Figure C.1. Sample Template for a "User Report" Prompt



There is a lot to be understood about individual systems from the overall ratings reported by users; however, the usefulness of the information is enhanced when we can distinguish among user groups under different circumstances. For example, a system that rates poorly on "easy to use" for all user groups can be assessed as having significant design flaws, while comparing user reports before and after the rollout of a new feature exposes whether that feature helped, hindered, or did not affect the usability of the system or application. A system that rates poorly on "easy to use" for only some models of end-user device may have a compatibility issue with those devices, while a poor rating for only some organizations may indicate a training issue. The selectable data fields that should be maintained for user reports include date of report, the end-user device model, user location and organization, and user service, rank or grade, and specialty (such as military occupation specialty for the Army, Air Force specialty code, or civilian job series). User service, organization, specialty, and rank or grade (or range of ranks or grades) can be used to approximate a user's role in interacting with a system. To illustrate, if the system being assessed is a civilian timekeeping system, DoD needs to be able to separately consider user reports from civilians who enter their time in the system and specialists who use the data for human resources functions.

Selectable fields for the user reports should be consistent with similar fields used in the technical performance parameter. For example, "location" could be a geographic location or a network address. It is also important that the user be asked for their input on only the two statements that comprise the

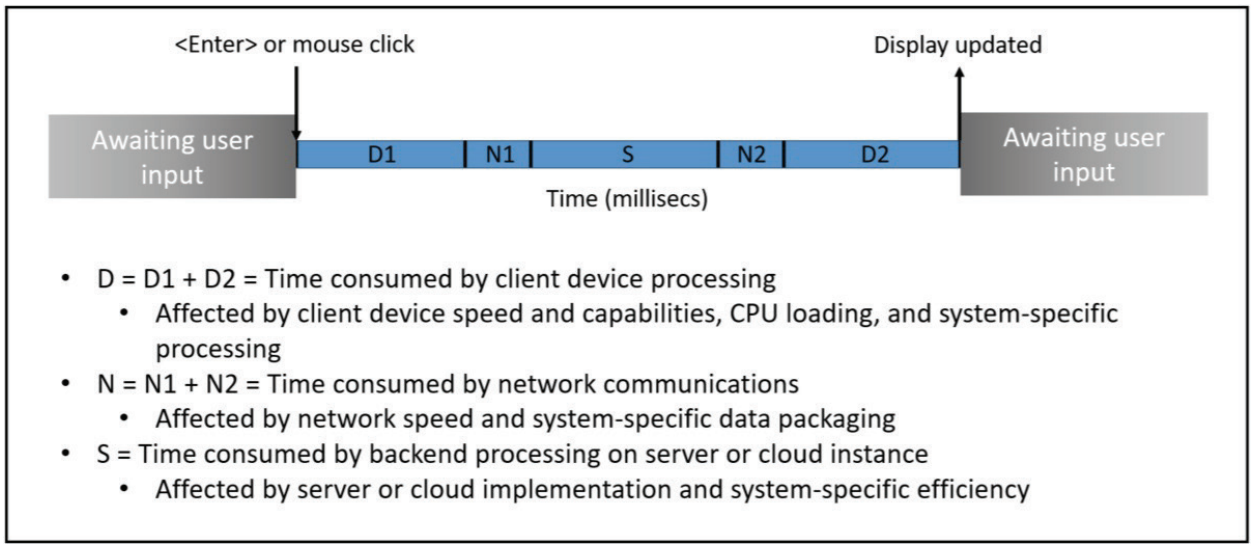
UMUX-Lite. Other data, such as location and organization, should be automatically associated with the entry.

Through a technical governance function, DoD can adjust the frequency with which any user is asked to report on any system, which selectable fields are necessary to distinguish user groups, and more. To prevent survey fatigue while collecting useful data, industry best practices suggest different frequencies to prompt for user input based on how often a user interacts with a system. For very frequent interactions (e.g., with Microsoft Office 365 products), a user report could be requested every three months. For very infrequently accessed applications, a user report could be requested after the user has interacted with the system two or three times. Another industry best practice is to request user reports after a new feature is deployed.[1]

Technical Performance

The second "value" metric, technical performance, is based on data measured at the end-user device. The time from when a user clicks a mouse button or presses <enter> until a new screen is displayed is the response time of the application. Unless the application runs entirely on the end-user device, that response time can be considered to involve three elements: processing on the end-user device, network processing and transit, and processing on the server, cloud, or other service provider equipment. For the technical performance parameter, those three elements are individually scored, and a weighted composite score is determined. Analyzing these scores for user groups under different

Figure C.2. Internal Elements of Technical Performance



NOTE: CPU = central processing unit.

[1] Jack Davies, "Think You're Sending Too Many Surveys? How to Avoid Survey Fatigue," *Qualtrics*, blog, 2019; Anna Kaley, "User-Feedback Requests: 5 Guidelines," Nielsen Norman Group, 2023; Manisha Khandelwal, "The Importance of Survey Frequency for Effective Feedback Strategies," *Survey Sensum*, blog, January 8, 2024.

conditions can indicate the source and degree of technological friction imposed on the users, stemming from either the system with which the scores are associated or the information environment external to that system.

The technical performance parameter consists of one score each for the end-user device (D), the network (N), and the server or cloud capability (S) processing time, and a weighted composite total (T) per transaction monitored at the end-user device. Each sub-score and the composite total are recorded on a scale of 0 to 100. This parameter requires monitoring software operating on a suitable subset of end-user devices, with data forwarded to a collector for DoD or responsive organizations.

The operation of any system can involve a vast number of individual transactions, potentially producing an intractable amount of technical performance data. For this reason, the technical governance function managing the framework must carefully choose which system activities to monitor, what metadata to associate with the transaction data, and the percentage of end-user devices to monitor. We make recommendations below for what we believe to be the minimum data necessary to effectively signal source and degree of friction affecting users, but this will have to be tuned throughout operation of the framework.

Determining a sub-score (D, N, or S) is a three-step process. The first step is a validity check, in which invalid transaction records are discarded (e.g., a transaction that records zero milliseconds for end-user device processing). In the second step, the time consumed in the response activity (end-user device, network, or server processing) is compared with human-relevant limits. Physiological studies show that the time it takes for humans to see and recognize an object ranges from 200 to 300 milliseconds on average and that deciding on a response can add as much as 500 milliseconds.[2] In his 1993 book *Usability Engineering*, Jakob Nielsen states as a rule of thumb that a response time of less than 100 milliseconds will seem instantaneous and that a response time of 1 second or less enables users to keep their train of thought. However, when experiencing a response time of 10 seconds or more, the user will likely start working on more than one task at a time, switching between windows and losing efficiency in doing so.[3] With these factors in mind, the framework is initially configured to score any element (D, N, or S) that takes less than 100 milliseconds at 100, and D and S elements that take more than 10 seconds or N that takes more than 5 seconds at 0.

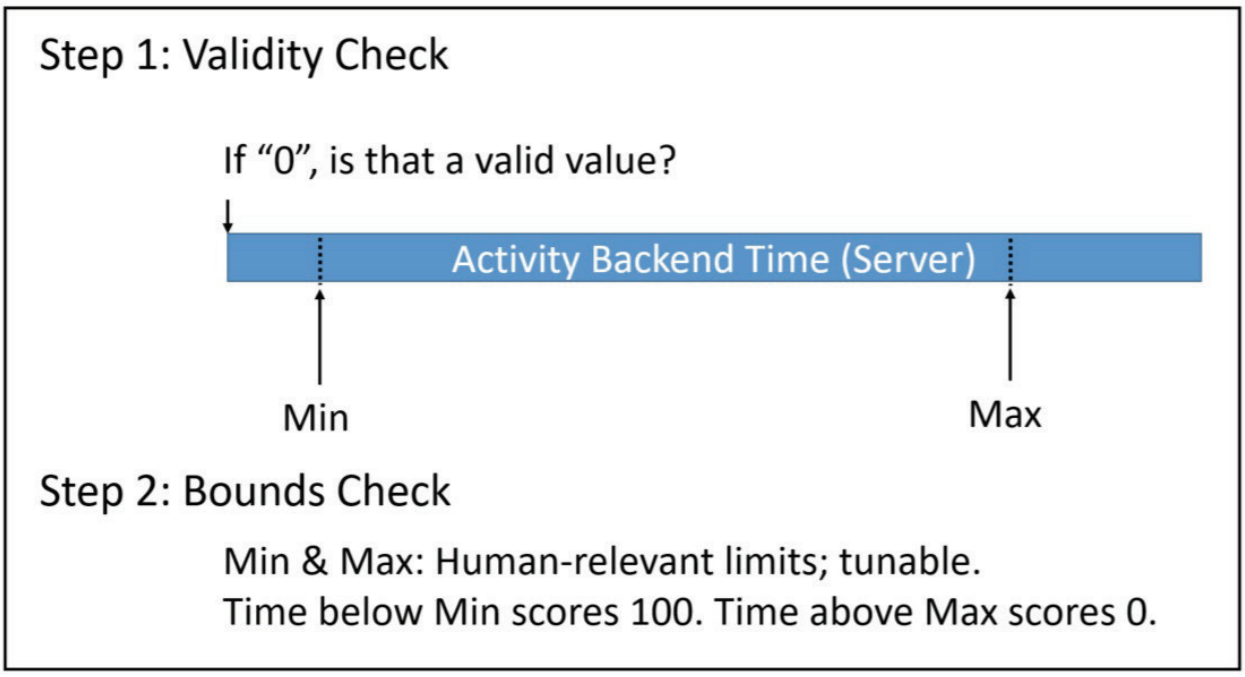
The third step in determining a transaction sub-score (if the sub-score was not determined in step 2) compares the time consumed at D, N, or S with the expected value for that type of activity on that system or application (e.g., an "Open Mail" activity on Microsoft Outlook). The "expected value" can be mean or median time for that activity determined through a burn-in period or a testing or industry benchmark. The time consumed in D, N, or S is scored on a linear scale from 0 to 100 such that if it matches the expected value it scores a 50 and at twice the expected value it scores 0.

---

[2] David L. Woods, John M. Wyma, E. William Yund, Timothy J. Herron, and Bruce Reed, "Factors Influencing the Latency of Simple Reaction Time," *Frontiers in Human Neuroscience*, Vol. 9, 2015; Richard P. Heitz, "The Speed-Accuracy Tradeoff: History, Physiology, Methodology, and Behavior," *Frontiers in Neuroscience*, Vol. 8, 2014.

[3] Jakob Nielsen, *Usability Engineering*, Morgan Kaufmann, 1994; S. K. Card, G. G. Robertson, and J. D. Mackinlay, "The Information Visualizer: An Information Workspace," CHI '91: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, La., April 28–May 2, 1991; Miller, 1968.

Figure C.3. Calculating Technical Performance: (a) Validity and Human-Relevant Boundary Checks; (b) Detailed Scoring per Element



(a)

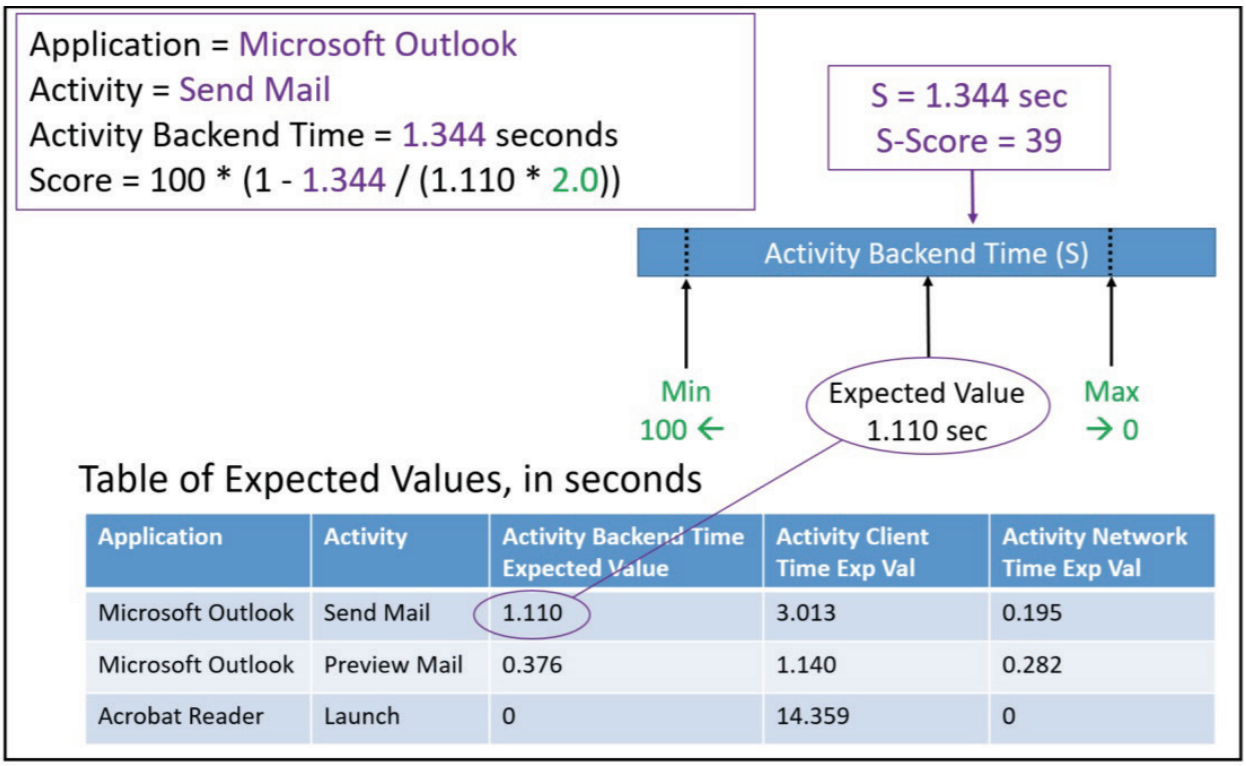


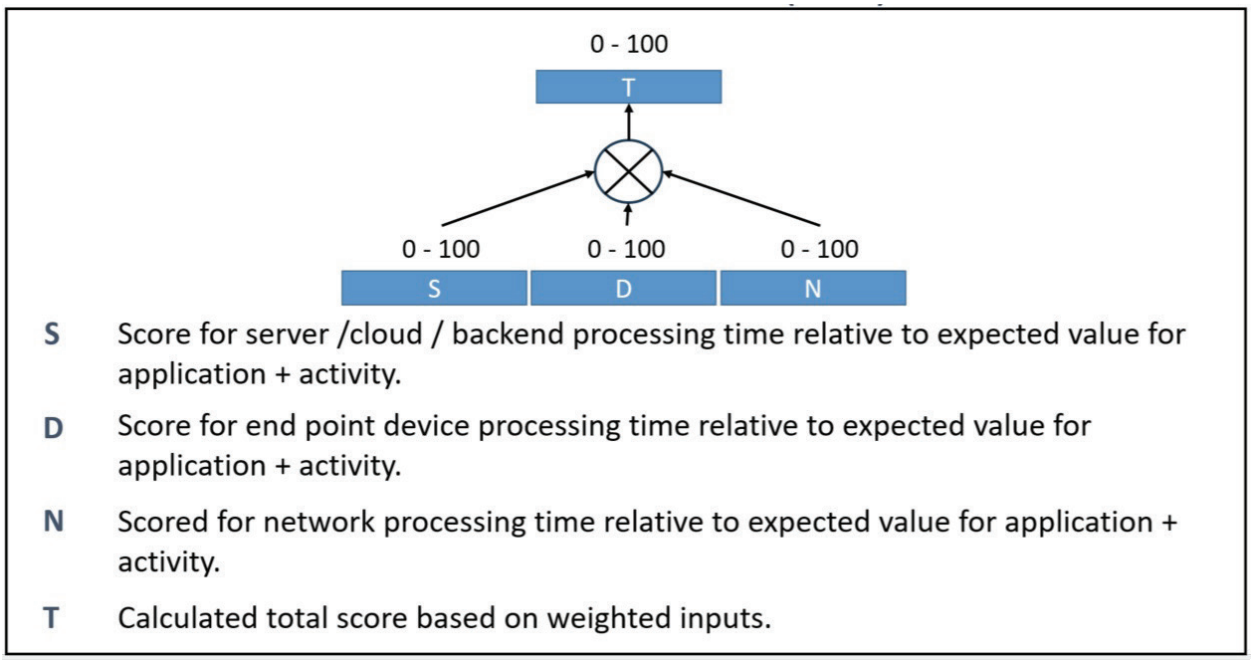
Table of Expected Values, in seconds

Application	Activity	Activity Backend Time Expected Value	Activity Client Time Exp Val	Activity Network Time Exp Val
Microsoft Outlook	Send Mail	1.110	3.013	0.195
Microsoft Outlook	Preview Mail	0.376	1.140	0.282
Acrobat Reader	Launch	0	14.359	0

(b)

The total score (T) is a weighted composite of D, N, and S that can be calculated when needed. "T" is intended to reflect the overall burden to the user posed by the combined friction of all the elements of an activity response. Weighting will have to be tuned until the overall score matches the assessment of burden by the technical governance entity. Initial framework configuration sets those weights at 70 percent for the D score, 10 percent for the N score, and 20 percent for the S score.

Figure C.4. Calculating Total Technical Performance Score



The technical performance parameter is intended to signal the source and degree of issues, not to troubleshoot those issues. Because of the potential for intractable volumes of data, the technical governance body should select the minimum data to accomplish that function. The value in scoring a transaction's network performance is to detect instances in which network function affects use of an application. To get that signal, the N scores for the same application and activity can be compared for user groups at different locations or at the same location over time. In this case, "location" should be a single parameter that carries the appropriate amount of data to detect that a network problem exists but does not need to isolate the exact problem. That task is left to network providers or maintainers. For that reason, "location" can be a geographical location, such as a base, post, camp, or station, or it can be a network location defined by a subnet. The technical governance body should determine which method of identifying a location best suits their purposes. Similarly, to identify issues stemming from capabilities of different end-user devices, the technical governance body should select one or two data elements that best describe a device, such as model number.

The Portfolio/Monetary Axis

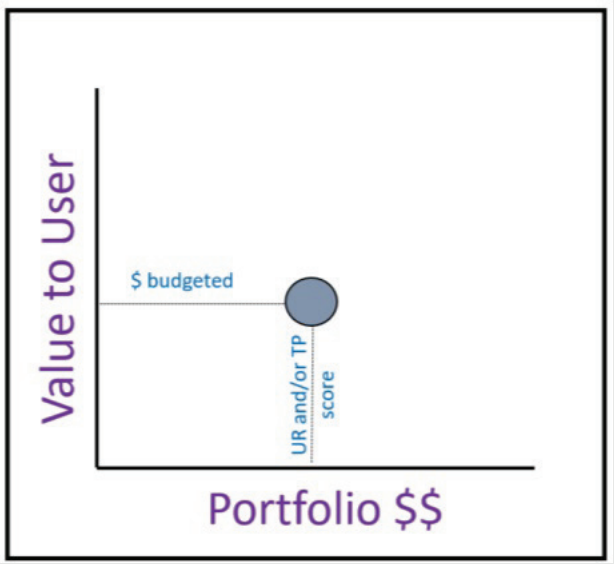
Though DoD is unable to accurately measure the amounts actually spent on IT and software for reasons stated earlier, we recommend using the yearly congressional budget appropriations to indicate the past and current year cost of each program element code.[4]

Thus, the horizontal axis of our proposed framework (see Figure C.5) shows the annual funding budgeted for the system(s) of interest from authoritative data sources. We recommend implementing a method to allow analysts to examine individual systems as well as a collection of systems related through the taxonomies of the DoD enterprise architecture, as used in IT portfolio management.

User Groups

Basic operation of the framework involves selecting a system or systems of interest and one or more user groups of interest and plotting the value score from that group's use of the system(s) against the budgeted annual cost of the system(s) in the framework as shown in Figure C.5.

Figure C.5. Top-Level Framework View



NOTE: UR = user report; TP = technical performance. Notional data used for illustration purposes.

User groups can be designated through the associations maintained with the user reports and technical performance data for that system, allowing selection of users by geographical or network location, end-user device, their organization, service, and date and time of the scoring event, or any combination of the above.

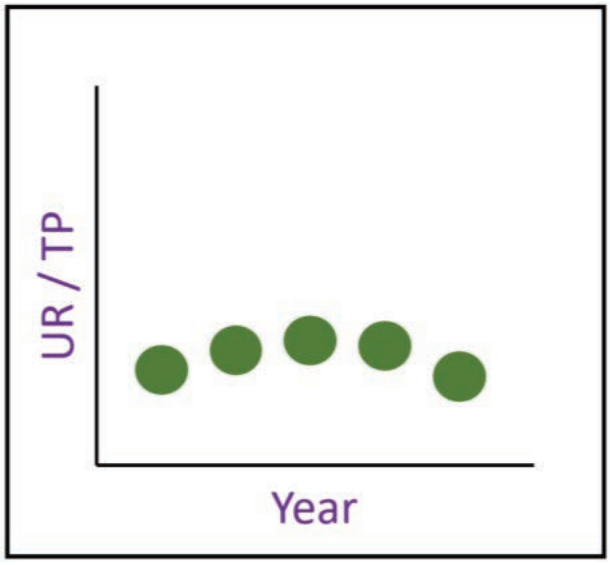
[4] Program element codes refer to a system or collection of similar systems and are the basis of all congressional appropriations to DoD.

Operations on Data

By comparing user value scores for the same system over time, for the same system under different IT environmental conditions, or for related systems, we can get valuable insights into the health of the system, impacts of IT environmental conditions, impact of investments, and system and interface design.

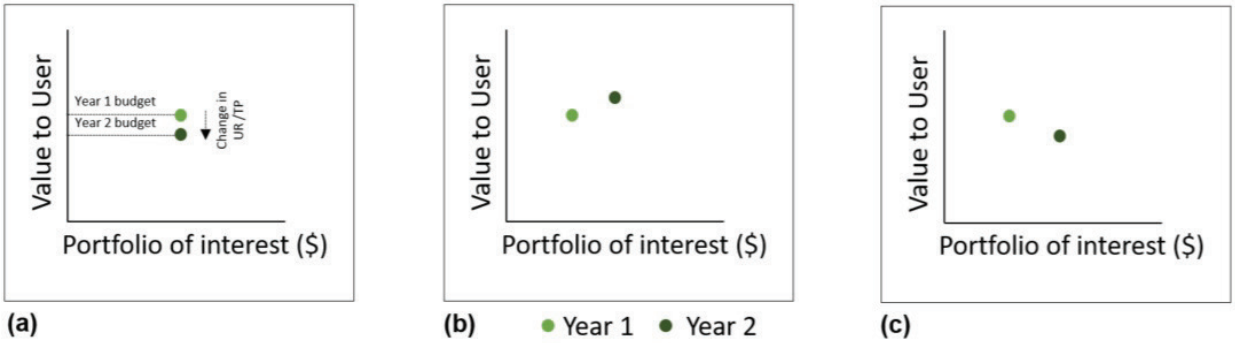
A simple view of user report and/or technical performance data over time provides feedback on a system's usability and usefulness, as shown in Figure C.6. When combined with funding data, the framework provides some information about the impact of investment decisions, as shown in Figure C.7.

Figure C.6. Changes in Value Ratings over Time for a Single System



NOTE: UR = user report; TP = technical performance. Notional data used for illustration purposes.

Figure C.7. Retrospective Analyses of Impact of Investment on Value to User: (a) Value Drops Under Equal Investment; (b) Value Increases with Increased Investment; (c) Value Drops After Increased Investment



NOTE: UR = user report; TP = technical performance. Notional data used for illustration purposes.

Figures C.7a through C.7c show a retrospective view of the impact of investment in a system or sub-portfolio (using notional data). In these examples, the same system or sub-portfolio is compared over two years. In Figure C.7a, the amount budgeted for Year 2 is the same as the amount budgeted for Year 1, but the user value rating dropped, which signals a problem. (The problem may come from within or outside the system or sub-portfolio of interest, and that follow-on analysis will be discussed in the next paragraph.) In Figure C.7b, additional investment resulted in improvement in value to the users, a positive sign. In Figure C.7c, however, additional investment resulted in a drop in the value to the user, which may indicate release of a poorly designed feature or that the network and/or end-user device is overloaded.

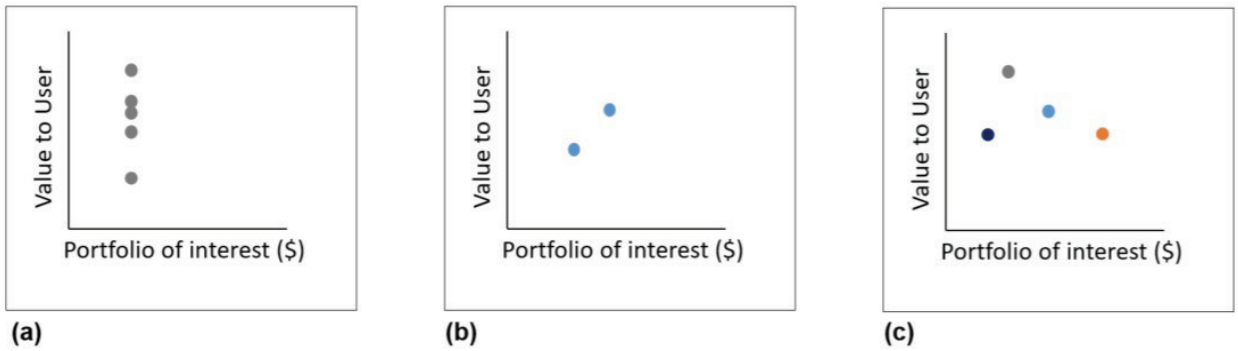
Although we caution that our framework is not meant to substitute for the more capable diagnostic tools maintained by the services, it may in some cases be able to provide insight into the source of user problems. Problems might come from within the system itself (e.g., a failing server farm or a change in mission demands to which the system did not adapt), outside the system or sub-portfolio (e.g., increased network or end-user device loading or failure), or a combination of internal and external factors. To test for these sources of problems, follow-on analysis could examine whether the value decrease is correlated to end-user device type, network location, or date, encompassing a period of time before and after the rollout of a new feature or a change in the IT environment. The user value rating can be decomposed to examine technical scores separately from user report scores and to examine sub-elements of user reports and technical performance independently to provide insight into the source and degree of the problems users encounter.

Figures C.8a through C.8c show a prospective analysis of whether further investment in a system or sub-portfolio is likely to positively affect value to the user and to what degree. Because of the heterogeneity of IT environments across the breadth of DoD, different user groups experience various IT services. Some geographic locations are served by newer and faster networking capabilities than others. Each service and agency conducts cybersecurity functions in different ways, often using different tools. Even when a tool is put in place for common use throughout DoD, there is usually a rollout process during which some users will be in an environment that includes the tool while others are not. Thus, by comparing technical performance and/or user report ratings for user groups that vary in these important ways, capability planners can get an indication of the likely user impact of the investment they are considering.

In Figure C.8a, users of the same system are divided into user groups by the type of their end-user device in order to reveal the impact of that end-user device on usability of the system. In Figure C.8b, value ratings for a group of users under special or experimental conditions (e.g., DAF enterprise IT as a service bases) are compared with standard users. In Figure C.8c, different service or agency approaches to the same or similar problem are compared in terms of cost and value (e.g., each service's financial accounting system).

Figure C.9 shows another way to view the data, which is to graph user report scores as a function of technical performance for a given application in different computing environments. In our analysis of the survey results, we noted that poor technical performance was correlated to low user satisfaction scores (i.e., the data points roughly lie along the diagonal axis as is seen for the gray and green dots in the figure), but that was not always true. For some applications in some environments (as seen in the red dot in the figure), the user experience is rated poorly despite having good technical performance;

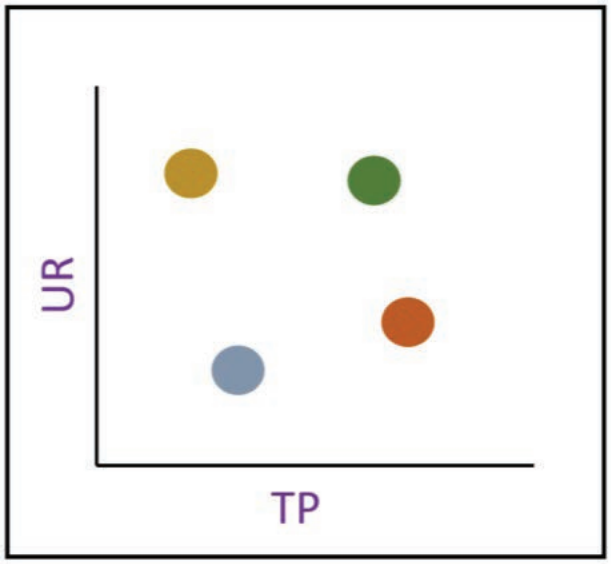
Figure C.8. Prospective Analyses of Impact of Investment on Value to User: (a) Value of a Single System Used on Different End-User Devices; (b) Value Functions of Two IT Service Delivery Modes; (c) Value Functions of Similar Systems in Different Organizations



NOTE: Notional data used for illustration purposes.

this is a good indicator of poor application design. In other cases (such as the brown dot in the figure), a highly rated user experience is achieved in spite of poor technical performance; this is a good indicator that the computing environment associated with the collection of this data point should be improved.

Figure C.9. Contrasting User Report Rating with Technical Performance Score for a System



NOTE: UR = user report; TP = technical performance. Notional data used for illustration purposes.

Appendix D

Lessons Learned, Insights, and Challenges Associated with the Military Health System's Electronic Health Record (MHS Genesis)

As noted in the review of the survey results (Appendix B), MHS Genesis is one of the five most often rated mission-critical applications in our survey. Given that the system had only recently completed its deployment to the full workforce and that a significant number of respondents were dissatisfied with it, there is an opportunity to better understand issues related to the rollout of a very complex, mission-critical application. Understanding that process from the point of view of various stakeholders provided a good test case as we formulated our recommendations. In particular, we used this test case to better understand the role of an ombudsman versus program office and mission-area leadership (see seventh recommendation). This appendix summarizes our thoughts.

Background

The Health.mil community, established in 2013 with the formation of the Defense Health Agency, consists of nearly 130,000 civilians and military personnel globally. This community is essential in supporting the agency's combat support mission.

In 2015, to enhance health care delivery, the Military Health System (MHS) introduced MHS Genesis, which is a unified electronic health record system for DoD and the Department of Veterans Affairs. The system, which reached its final clinic on March 9, 2024, aims to improve data sharing and interoperability.

However, despite its ambitious goals, MHS Genesis has faced significant challenges. The Health.mil civilian population in our survey had a strong response rate—indicating they were motivated to report issues. Of the survey responses that mentioned MHS Genesis, about one-third (35 percent) came from respondents inside Health.mil. The distribution of respondents to our user-satisfaction questions for MHS Genesis is shown in Figure B.5 and indicates significant dissatisfaction.

The level of dissatisfaction we found in a recently deployed business system (a not uncommon occurrence) underscores the need for continuous improvement and effective problem-solving during the deployment of a complex IT system. By examining the issues identified with MHS Genesis and the processes the team used in deploying the system (many of which are exemplary), we aim to

highlight lessons learned, insights, and challenges that can guide the ongoing improvement of software and IT systems within the business mission area. Key questions include the following:

- What does it take to put medical IT and software systems on a path toward continual learning and improvement, making these systems more useful and responsive?
- In what ways can software integration processes be improved?
- What might DoD consider in implementing an ombudsman approach to identify and communicate problems?

The following section outlines our approach for this initial examination of these questions.

Approach

We document the lessons learned, insights, and challenges associated with MHS Genesis, drawing from several sources, including

- other research regarding the deployment of MHS Genesis[1]
- interviews with the current Defense Health Agency CIO and the outgoing Defense Health Agency Chief Health Informatics Officer (as of July 2024)
- over 4,500 survey responses from armed forces, documenting challenges with software and IT
- initial review of an ombudsman approach for DoD
- initial review of organizational policies and resources regarding data sharing, sensitivity, security, and foresight approaches in vendor selection to improve future software integration efforts.

It is important to note that we do not provide an exhaustive list of lessons learned, challenges, or issues. Rather, the discussion that follows serves as a starting point for future software integration efforts across DoD. More research is needed, particularly regarding the ombudsman approach, to fully understand and address the complexities involved—including resource and funding constraints, effective implementation within DoD's organizational structure, and DoD's unique operational dynamics related to data sharing, policies, and security.

Key Lessons Learned, Insights, and Challenges

Invest in Proactive Monitoring to Enhance Situational Awareness and Mitigate Potential Risks Effectively

Proactive monitoring of IT and software issues helps organizations anticipate and resolve them before they become critical. Automated monitoring tools can be used to continuously track the performance of IT and software systems in terms of network traffic, system logs, usage patterns, and the like. These tools can detect outage or performance issues in real time before severely disrupting

[1] Government Accountability Office, *Electronic Health Records: DoD Has Deployed New System but Challenges Remain*, GAO-24-106187, April 18, 2024;

workflow and affecting users. Frequent audits of IT infrastructure can also help to ensure systems are functioning optimally. Audits are important for identifying areas needing improvement, and these areas can be attended to using a prioritization scheme. Improvements and resolutions to issues depend on understanding the root cause of the issue, which is typically a function of how user workflow and technology interact. Standardization of workflow is therefore especially helpful for developing a baseline understanding of users' experiences, the issues that they encounter, and the solutions needed.

The benefits of proactive IT and software monitoring also depends on whether the workforce capacity is sufficient to respond to and resolve issues in a timely manner. In anticipation of issues during the initial implementation (or rollout) of a new system, the number of specialists supporting the process can be plentiful and include personnel such as trainers, issue resolution specialists, and account coordinators. Importantly, sustaining and potentially growing the workforce capacity beyond initial implementation should be planned for, especially for systems that are deployed at scale and take time for users to build competency.

Develop or Optimize Channels for User Feedback to Improve Engagement and Drive Continuous Improvement

Beyond using automated monitoring to trace IT and software issues, users can provide valuable feedback regarding the shortfalls of a new system. As with automated monitoring, gathering regular and consistent feedback provides important insight into users' experiences. This feedback is possible once a new system has been introduced across a sufficient number of user groups who have had time to learn and adapt to the new processes. Feedback can be collected through approaches such as surveys and interviews, as well as through informal documentation of shared experiences. Further, committees that represent different types of user groups can be established to provide focused feedback and explore potential improvements. Such committees should include leading domain experts who are dedicated to identifying and prioritizing the most important problems to solve. Finally, for users who prefer to report problems privately, ombudsman programs can provide an accessible way for users to share their IT and software issues with confidentially.

Considerations for Implementing an Ombudsman Approach

Organizations of various types and sizes often have an ombudsman office in place to help their personnel, members, customers, or other individuals affiliated with the organization resolve issues—usually interpersonal in nature. Individuals have the option to seek assistance from an ombud when

- they believe they have been treated unfairly
- they do not know whom to address their concern to
- they wish to better understand policies
- they want an independent third-party to help them resolve a problem.[2]

[2] U.S. Department of State, "Office of the Ombuds," webpage, undated.

According to the International Ombuds Association, an ombudsman must be an independent, impartial, confidential, and informal neutral mediator.[3] They must operate outside of the standard organizational hierarchy to be effective, which also means they have no incentive to prevent issues from reaching senior leaders.

Although this role is typically designated for interpersonal or otherwise general grievances an employee might have, there may be opportunities to utilize an ombudsman office to capture and address IT issues within and across DoD, such as those uncovered for MHS Genesis. The ombudsman could incorporate information about a common issue from disparate sources (in the case of MHS Genesis, the sources were physicians, nurses, dentists, and the like), and compile reports to create a unified picture of the problems that are occurring and reoccurring. An ombudsman office could offer a distinct type of assistance beyond traditional IT help-desk responsibilities. Unlike a standard IT help desk, the ombudsman office would not focus on resolving technical issues but would instead evaluate and address *how well* the organization manages IT-related concerns. Some potential questions and issues an IT-focused ombud could address include the following:

- What standard, routine IT-related procedures are not working properly for personnel?
- Is there a consistent delay in personnel receiving assistance or in having their IT issues resolved?
- What have been the most frequent complaints in the past year, quarter, or several weeks?
- Are there any trends or systematic issues that persist for personnel using IT systems?

Some important considerations for DoD in standing up one or several additional ombudsman offices include whether it would be feasible to have existing ombudsmen absorb (or be encouraged to absorb) IT-specific issues from personnel.[4] Alternatively, there could be a dedicated "IT-ombud" for each DoD vocational community, such as clinicians, educators, or lawyers.

To establish an effective IT-centric ombudsman program, it is crucial to define clear objectives. These objectives should focus on improving user experience and enhancing employee efficiency. By setting specific goals, the program can ensure that its efforts are aligned with the DoD business mission areas.

Next, it is essential to implement robust reporting and feedback mechanisms. These systems will capture user issues efficiently and create a feedback loop, allowing senior personnel to act on behalf of users. This continuous cycle of feedback and improvement will help address user concerns promptly and effectively.

Standardizing procedures for personnel interaction with the IT-ombudsman is another critical step. Clear and well-documented procedures should be developed to guide employees on how to seek assistance. Ensuring these procedures are easily accessible and well communicated will help streamline the process and make it more user-friendly.

[3] International Ombuds Association, "Welcome to the Ombuds Toolkit," webpage, undated.

[4] Parts of DoD that already have ombudsman offices include the Office of the Inspector General and the Washington Headquarters Service. For more information on these offices, see Alison Whaley, "Ombuds," DoD Office of Inspector General, webpage, undated; Washington Headquarters Service, "WHS" Ombudsman Office," webpage, undated.

In addition, designing a detailed workflow for reporting and resolving issues is vital. This workflow should outline each step from the initial report to the final resolution, thereby ensuring transparency and accountability throughout the process. A well-defined workflow will help in managing issues systematically and efficiently.

In the DoD context, establishing a chain of command for the IT ombudsman is necessary. This structure should operate outside the standard hierarchy to effectively manage critical business mission areas while maintaining alignment with overall organizational goals. By defining a clear chain of command, the program can ensure that it operates smoothly and effectively within the unique DoD environment.

Invest in Continual Learning and Improvement Processes to Drive Sustained Growth and Innovation

To enhance the capabilities and benefits of new IT and software systems, learning and improving should be ongoing goals. A unification of workflow that results in standardized systems and processes is essential for learning and improvement, especially during the early implementation of a system. In this unified workflow, the users, their processes, and technology should be integrated fully. A standardized and integrated system supports training efforts, such that a pool of experts can be created to help train other users and troubleshoot their issues. For example, programs can be developed to leverage experienced users as others transition to the new system. This approach is effective, because expert users are knowledgeable about both the system and the functional area for which it is designed. These expert users can help train their peers within the working environment for which the system is intended (e.g., doctors supporting doctors in the case of MHS Genesis). Leaning on communities of practice in this way provides opportunities for continuous learning, which is essential for building competency. Establishing competency starts with a minimal required understanding of a system and grows over time. If too many variations of a system exist initially, it is less likely that a pool of experts will be available for peer-to-peer training and continual learning.

Learning opportunities are also an important aspect of change management. In addition to learning a new system, users need to be educated on the reasons for developing a new system and the benefits the system is expected to bring. Part of this education can include highlighting differences in processes between the legacy and new systems. In addition, senior leadership needs to drive governance for how a new system is implemented. This leadership is especially important when multiple organizations are involved in a large-scale and complex rollout. A strategic alignment strategy could be considered.[5] In this approach, the current project management practice of each organization is assessed, and areas of strength, weakness, and commonality are identified. Early on, organizations should also align cost management and budgetary tracking standards and emphasize a risk management plan. Finally, this approach recommends establishing a communication plan to enable transparent communication and feedback mechanisms across organizations.

[5] One such approach can be found in Sergey Filippov, Herman Mooi, Roelef van der Weg, and Laurent-Jan van der Westen, "Strategic Alignment of the Project Portfolio: An Empirical Investigation," PMI® Research and Education Conference, Limerick, Munster, Ireland. Newtown Square, PA: Project Management Institute, 2012.

Enhance Vendor Management and Prioritize Data Integration for Effective Change Management and Long-Term Success

Learning from past implementations of new systems and leveraging known successes can significantly guide change management strategies for new IT and software systems. By analyzing previous efforts, organizations can identify best practices and avoid repeating mistakes, thereby streamlining the implementation process. A critical lesson learned is that working toward a common data format offers the best opportunity to develop a single integrated system, as opposed to the fragmented systems exemplified by MHS Genesis. The lack of alignment on a common data format has led to inefficiencies, such as medical staff having to enter information twice—one for each system—resulting in potential lost working hours and reduced productivity.

Conducting a strategic review of each organization's vendor management policies, particularly in areas of selection and acquisition for IT and software systems, can be highly beneficial. This review can enhance understanding of the policies' technology foresight capabilities. Leadership can gain valuable insights through a robust technology foresight view embedded in IT vendor selection and acquisition processes. These insights can inform long-term investment decisions, which are crucial for effective change management.

Large-scale efforts involving the integration of IT and software systems should prioritize alignment on data integration techniques and strategies from the outset. During the requirements-gathering process, it is essential to consider various data integration methods to ensure a seamless transition to a common data format. This approach facilitates the progression toward a unified IT or software system. Key data integration techniques to consider include

- data consolidation: combining data from multiple sources into a single source
- data federation: allowing users to access data from multiple sources
- data transformation: converting data from one format to another
- data propagation: copying data from one location to another
- middleware data integration: integrating data from different sources
- data warehousing: creating a centralized repository
- manual data integration: manually entering data.

Identifying pain points in existing incident management systems is critical for optimizing how issues are tracked, categorized, and resolved. DoD should document and analyze any patterns in reported issues to enhance the efficiency and effectiveness of their incident management processes.

Improving IT vendor management policies to include a more uniform and robust technology review process is essential. This improvement should incorporate technology foresight capabilities during the vendor selection process. Utilizing data from sources such as Gartner reports or other technology business trend reports can provide valuable insights that inform better decisionmaking.

Summary

The successful implementation and management of IT and software systems such as MHS Genesis require a multifaceted approach that incorporates proactive monitoring, user feedback, and

robust change management strategies. Investing in proactive monitoring enhances situational awareness and mitigates potential risks by detecting issues in real time and addressing them before they escalate. This approach, coupled with frequent audits and a standardized workflow, ensures optimal system performance and user satisfaction.

Developing or optimizing channels for user feedback is equally important. By gathering regular and consistent feedback through surveys, interviews, and specialized committees, DoD can gain valuable insights into user experiences and identify areas for improvement. For DoD, establishing an IT-centric ombudsman program with clear objectives and robust reporting mechanisms can significantly enhance user experience and employee efficiency. Standardizing procedures for interaction with the IT ombudsman and designing a detailed workflow for issue resolution are critical steps in this process. In addition, defining a chain of command for the IT ombudsman within the DoD context ensures that the program operates smoothly and aligns with organizational goals.

Continual learning and improvement processes are essential for driving sustained growth and innovation. By standardizing systems and processes, organizations can create a pool of expert users who can provide peer-to-peer training and support. This approach fosters continuous learning and competency building, which are vital for the successful adoption of new systems.

Further, enhancing vendor management policies and prioritizing data integration strategies are crucial for effective change management. Learning from past implementations and conducting strategic reviews of vendor management policies can inform long-term investment decisions and streamline the implementation process. Aligning data integration techniques from the outset ensures a seamless transition to a unified IT system, optimizing incident management and decisionmaking.

By addressing these areas comprehensively, DoD can better navigate the complexities of modern IT landscapes, achieve its business mission-area objectives, and ensure long-term success.

Appendix E

Full Text of Fiscal Year 2023 National Defense Authorization Act, Section 241

Sec. 241 Study on Costs Associated with Underperforming Software and Information Technology.

(a) Study Required. The Secretary of Defense shall seek to enter into a contract or other agreement with an eligible entity to conduct an independent study on the challenges associated with the use of software and information technology in the Department of Defense, the effects of such challenges, and potential solutions to such challenges.

(b) Elements. The independent study conducted under subsection (a) shall include the following:

(1) A survey of members of each Armed Force under the jurisdiction of the Secretary of a military department to identify the most important software and information technology challenges that result in lost working hours, including

(A) An estimate of the number of working hours lost due to each challenge and the cost of such lost working hours;

(B) The effects of each challenge on service member and employee retention; and

(C) Any negative effects of each challenge on a mission of the Armed Force or military department concerned.

(2) A summary of the policy or technical challenges that limit the ability of each Secretary of a military department to implement needed software and information technology reforms, which shall be determined based on interviews conducted with individuals who serve as chief information officer (or equivalent position) in a military department.

(3) Development of a framework for assessing underperforming software and information technology, with emphasis on foundational information technology to standardize the measurement and comparison of programs across the Department of Defense and its component organizations. Such a framework shall enable the assessment of underperforming software and information technology based on:

(A) designs, interfaces, and functionality which prioritize user experience and efficacy;

(B) costs due to lost productivity;

(C) reliability and sustainability;

(D) comparisons between:

(i) outdated or outmoded information technologies, software, and applications; and

(ii) modern information technologies, software and applications;

(E) overhead cost for software and information technology in the Department compared to the overhead costs for comparable software and information technology in the private sector;

(F) comparison of the amounts the Department planned to expend on software and information technology versus the amounts actually spend for such software and services;
(G) the mean amount of time it takes to resolve technical problems reported by users;
(H) the average rate, expressed in time, for remediating or patching weaknesses or flaws in information technologies, software, and applications;
(I) workforce training time;
(J) customer satisfaction.

(4) The development of recommendations:
(A) To address the challenges identified under paragraph (1), and
(B) To improve the processes through which the Secretary provides software and information technology throughout the Department, including through:
(i) Business process re-engineering;
(ii) Improvement of procurement or sustainment processes;
(iii) Remediation of hardware and software technology gaps;
(iv) The development of more detailed and effective cost estimates.

(c) Report Required. Not later than one year after the date of the enactment of this Act, the eligible entity that conducts the study under subsection (a) shall submit to the Secretary of Defense and the congressional defense committees a report on the results of such study.

(d) Definition. In this section:
(1) The term "eligible entity" means an independent entity not under the direction or control of the Secretary of Defense, which may include a department or agency of the Federal Government outside of the Department of Defense.
(2) The term "software and information technology" does not include embedded software and information technology used for weapons systems.

Abbreviations

CIO	chief information officer
CONUS	continental United States
DAF	Department of the Air Force
DoD	Department of Defense
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DoN	Department of the Navy
FY	fiscal year
GAO	Government Accountability Office
IT	information technology
MHS	Military Health System
NDAA	National Defense Authorization Act
NDRI	National Defense Research Institute
OCONUS	outside the continental United States
OSD	Office of the Secretary of Defense
UMUX	Usability Metric for User Experience
VPN	virtual private network

References

Bangor, Aaron, Philip Kortim, and James Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *Journal of User Experience*, Vol. 4, No. 3, May 2009.

Boyatzis, R. E., *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage Publications, 1998.

Brooks, John, "SUS: A Retrospective," *Journal of User Experience*, Vol. 8, No. 2, February 2013.

Bur, Jessie, "Have the DoD's Special Hiring Practices Hurt More Than Helped?" *Federal Times*, May 6, 2021.

Card, S. K., G. G. Robertson, and J. D. Mackinlay, "The Information Visualizer: An Information Workspace," CHI '91: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, LA, April 28–May 2, 1991.

Commission on Planning, Programming, Budgeting, and Execution Reform, *Interim Report*, August 2023.

Congressional Research Service, *The Fundamentals of Military Readiness*, R46559, October 2, 2020.

Congressional Research Service, *Department of Defense Budget: An Orientation*, November 2021.

Congressional Research Service, *DoD Planning, Programming, Budgeting, and Execution (PPBE): Overview and Select Issues for Congress*, R47178, July 11, 2022.

Dabrowski, Jim, and Ethan Munson, "40 Years of Searching for the Best Computer System Response Time," *Interacting with Computers*, Vol. 23, No. 5, 2011.

Davies, Jack, "Think You're Sending Too Many Surveys? How to Avoid Survey Fatigue," *Qualtrics*, blog, June 25, 2019.

Decision Lab, "Why Do We Focus on Items or Information That Are More Prominent and Ignore Those That Are Not?," blog post, undated.

DeCuir-Gunby, Jessica T., Patricia L. Marshall, and Allison W. McCulloch, "Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project," *Field Methods*, Vol. 23, No. 2, 2011

Defense Business Board, *Recommendations to Improve IT User Experience Within DoD*, February 2, 2023.

Defense Innovation Board, *Software Is Never Done: Refactoring the Acquisition Code for Competitive Advantage*, May 3, 2019.

Defense Innovation Board, *Department of Defense Workforce: Competing for Digital Talent*, September 15, 2020.

Defense Manpower Data Center, *Number of Military and DoD Appropriated Fund Civilian Personnel*, dataset, September 30, 2023.

Department of Defense, *DoD Digital Modernization Strategy*, June 2019.

Department of Defense, *Department of Defense (DoD) Zero Trust Reference Architecture*, Version 2.0, July 2022.

Department of Defense, *Software Modernization Implementation Plan Summary*, March 29, 2023.

Department of Defense Chief, *Information Enterprise Architecture v3.0 Increment 2 Overview Document*, July 22, 2023, Not available to the general public.

Department of Defense Directive 8115.01, *Information Technology Portfolio Management*, October 10, 2005.

Department of Defense Directive 7045.20, *Capability Portfolio Management*, September 25, 2023.

Department of Defense Instruction 8115.02, *Information Technology Portfolio Management Implementation*, October 30, 2006.

Department of Defense Instruction 5000.75, *Business Systems Requirements and Acquisition*, Change 2, January 2020.

Department of Defense, Office of the Chief Information Officer, *Department of Defense Tech Debt Metrics Report*, June 2023.

Department of the Navy, "Strategic Intent to Implement World-Class Alignment Metrics," draft memo for DoN CIO signature, pre-decisional, undated.

DoD—*See* Department of Defense.

DoDD—*See* Department of Defense Directive.

DoDI—*See* Department of Defense Instruction.

DoN—*See* Department of the Navy.

Filippov, Sergey, Herman Mooi, Roelef van der Weg, and Laurent-Jan van der Westen, "Strategic Alignment of the Project Portfolio: An Empirical Investigation," PMI® Research and Education Conference, Limerick, Munster, Ireland, Newtown Square, PA: Project Management Institute, 2012.

GAO—*See* Government Accountability Office.

Gary B, "App Crash Research: Is Your App Below or Above the Benchmark?" *Medium*, August 17, 2020.

Government Accountability Office, *DoD Space Acquisitions: Including Users Early and Often in Software Development Could Benefit Programs*, GAO-19-136, March 18, 2019.

Government Accountability Office, *DoD Software Acquisition: Status of and Challenges Related to Reform Efforts*, GAO-21-105398, September 30, 2021.

Government Accountability Office, *Defense Infrastructure: DoD Should Better Manage Risks Posed by Deferred Facility Maintenance*, GAO-22-104481, January 31, 2022.

Government Accountability Office, *Leading Practices: Agency Acquisition Policies Could Better Implement Key Product Development Principles*, GAO-22-104513, March 10, 2022.

Government Accountability Office, *IT Systems Annual Assessment: DoD Needs to Improve Performance Reporting and Development Planning*, GAO-23-106117, June 2023.

Government Accountability Office, *Electronic Health Records: DoD Has Deployed New System but Challenges Remain*, GAO-24-106187, April 18, 2024.

Hagel, Chuck, Secretary of Defense, "Guidance for Implementation of Pay Our Military Act," Memorandum for Components and Defense Agencies, Department of Defense, October 5, 2013. As of September 2024: https://cdn.govexec.com/media/gbc/docs/pdfs_edit/poma_implementation_guidance.pdf

Harper, Jon, "Air Force Grappling with Budgetary Implications of Enterprise IT as a Service," *DefenseScoop*, July 21, 2023.

Heitz, Richard P., "The Speed-Accuracy Tradeoff: History, Physiology, Methodology, and Behavior," *Frontiers in Neuroscience*, Vol. 8, 2014.

International Ombuds Association, "Welcome to the Ombuds Toolkit," webpage, undated. As of December 15, 2024:
https://www.ombudsassociation.org/assets/docs/docs_2022/IOA%20External%20Audience%20Toolkit%20.pdf

Joint Staff, Chairman of the Joint Chiefs of Staff Instruction 5123.01, *Charter of the Joint Requirements Oversight Council and Implementation of the Joint Capabilities Integration and Development System, Enclosure D*, October 30, 2021.

Joint Staff, *Terminology Repository of DOD Issuances*, Version 14, September 15, 2023.

Kaley, Anna, "User-Feedback Requests: 5 Guidelines," Nielsen Norman Group, 2023.

Kanaan, Michael J., "An Open Letter, Fix Our Computers!," X.com, January 2022.

Khandelwal, Manisha, "The Importance of Survey Frequency for Effective Feedback Strategies," *Survey Sensum*, blog, January 8, 2024.

Legris, Paul, John Ingham, and Pierre Collerette, "Why Do People Use Information Technology? A Critical Review of the Technology Acceptance Model," *Information & Management*, Vol. 40, No. 3, 2003, pp. 191–204.

Lewis, James R., Brian S. Utesch, and Deborah E. Maher, "Investigating the Correspondence between UMUX-Lite and SUS Scores," in A. Marcus, ed., *Design, User Experience, and Usability: Design Discourse*, Springer, 2015.

Maslach, Christina, and Susan E. Jackson, "The Measurement of Experienced Burnout," *Journal of Occupational Behaviour*, Vol. 2, 1981.

McKeown, David W., DoD Senior Information Security Officer, *Continuous Authority to Operate*, OSD Memorandum, February 4, 2022.

Miles, Matthew B., A. Michael Huberman, and Johnny Saldaña, *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed., Sage Publications, 2014.

Miller, Jason, "Navy Used Threat of Cyber Vulnerability to Expand VDI," *Federal News Network*, February 16, 2024.

Miller, Robert, "Response Time in Man-Computer Conversational Transactions," *Proceeding of the AFIPS Fall Joint Computer Conference*, Vol. 33, p1968.

Mills, Patrick, Muharrem Mane, Kenneth Kuhn, Anu Narayanan, James D. Powers, Peter Buryk, Jeremy M. Eckhause, John G. Drew, Kristin F. Lynch, *Articulating the Effects of Infrastructure Resourcing on Air Force Missions: Competing Approaches to Inform the Planning, Programming, Budgeting, and Execution System*, RAND Corporation, RR-1578-AF, 2017. As of January 2024:
https://www.rand.org/pubs/research_reports/RR1578.html

Moore, James W., "What Is the Sense of Agency and Why Does It Matter?" *Frontiers in Psychology*, Vol. 7, August 29, 2016.

National Institute of Standards and Technology, *Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy*, NIST SP 800-37 Rev 2, December 2018.

Nielsen, Jakob, *Usability Engineering*. Morgan Kaufmann, 1994.

Nord, Robert, and Ipek Ozkaya, "10 Years of Research in Technical Debt and an Agenda for the Future," *SEI Blog*, August 22, 2022.

Office of Personnel Management, "Hiring Process Analysis Tool," Human Capital Management/ Hiring Reform, webpage, undated. As of January 2024: https://www.opm.gov/policy-data-oversight/human-capital-management/hiring-reform/hiring-process-analysis-tool/

Office of Personnel Management, "Government-Wide Authority," Direct Hire Authority, website, undated. As of November 2024: https://www.opm.gov/policy-data-oversight/hiring-information/direct-hire-authority/#url=Governmentwide-Authority

Office of the Secretary of Defense, Comptroller, "FY 24 Budget Submittal, Exhibit M-1," undated-a.

Office of the Secretary of Defense, Comptroller, "FY 24 Budget Submittal, Exhibit O-1," undated-b.

Office of the Secretary of Defense, Director, Administration and Management, *Achieving "Mission Ready": How OSD's IT Enterprise Can Benefit from Refreshed Strategy, Leadership, and Resourcing*, July 2022.

Office of the Secretary of Defense, Director, Administration and Management, *OSD IT Enterprise Implementation Plan: The Initial Steps of the Journey Toward Improved Digital Experience*, February 2023.

OSD—*See* Office of the Secretary of Defense.

PEO Digital, "PEO World-Class Alignment Metrics 101," undated presentation.

PEO Digital, "World-Class Alignment Metrics," presentation, 2023.

Public Law 117–263, James M. Inhofe National Defense Authorization Act for Fiscal Year 2023, December 23, 2022.

Rathbun, Jane, "IT Pilots Show Promise Toward 'Fix My Computer,'" Department of the Navy, Chief Information Officer, webpage, Aug 1, 2023. As of January 2024: https://www.doncio.navy.mil/ContentView.aspx?id=16360

Robson, Sean, Bonnie L. Triezenberg, Samantha E. DiNicola, Lindsey Polley, John S. Davis II, and Maria C. Lytell, *Software Acquisition Workforce Initiative for the Department of Defense: Initial Competency Development and Preparation for Validation*, RAND Corporation, RR-3145-OSD, 2020. As of January 3, 2024: https://www.rand.org/pubs/research_reports/RR3145.html

Serbu, Jared, "Navy Ready to Start Implementing Fixes to Notoriously Slow Computers," *Federal News Network*, May 19, 2023.

Serbu, Jared, and Scott Maucione, "Congress Taps Brakes on DoD Project to Reform IT Funding," *Federal News Network*, March 14, 2022.

thegirlisok, "Fix Our Computers!," archived post from the r/Military subreddit thread, undated. As of January 2024: https://www.reddit.com/r/Military/comments/sdlvk7/fix_our_computers/

Triezenberg, Bonnie L., Mary Lee, Kristen Van Abel, Arianne Collopy, Brian Dolan, Sandra Kay Evans, Marissa Herron, and Joshua Steier, *Essential Utilities: Developing an Investment Strategy for U.S. Space Force Mission-Enabling Infrastructure*, RAND Corporation, 2024, Not available to the general public.

Triezenberg, Bonnie L., Jason M. Ward, Jonathan Cham, Devon Hill, Sean Robson, and Jeff Fourman, *The Composition and Employment of Software Personnel in the U.S. Department of Defense: An Initial Analysis*, RAND Corporation, RR-A520-1, 2020. As of January 2024: https://www.rand.org/pubs/research_reports/RRA520-1.html

Turner, Darren, "Flank Speed: Exceptions to Policy," *PEO Digital News*, May 11, 2023.

Ursin, Jani, Katja Vähäsantanen, Lynn McAlpine, and Päivi Hökkä, "Emotionally Loaded Identity and Agency in Finnish Academic Work," *Journal of Further and Higher Education*, Vol. 44, No. 3, 2020.

U.S. Central Command, Assessment and Analysis Division, *Operation IRAQI FREEDOM—By the Numbers*, April 30, 2003.

U.S. Code, Title 10, Section 2223 (b). Information Technology: Additional Responsibilities of Chief Information Officer of Military Department.

U.S. Department of State, "Office of the Ombuds," webpage, undated. As of December 15, 2024: https://www.state.gov/resources-office-of-the-ombuds/

Walmart, "How Many People Work at Walmart?" webpage, undated. As of January 2024, https://corporate.walmart.com/askwalmart/how-many-people-work-at-walmart

Washington Headquarters Service, "WHS Ombudsman Office," webpage, undated. As of December 15, 2025: https://www.whs.mil/About-WHS/Directorates/Human-Resources-Directorate-HRD/Careers/WHS-Ombudsman-Office/

Whittall, Colt, "How We Are Fixing Our Computers," *Medium*, March 16, 2023a.

Whittall, Colt, "Update: How We Are Fixing Our Computers," presentation, August 28, 2023b.

Whittall, Colt, "10/24/24 Colt Whittall Comments re Rand Report 'Underperforming Software and Information Technology in DOD,'" email to Sarah Zabel, November 6, 2024.

Wilson, Bradley, Padmaja Vedula, Aimee Bower, Timothy Parker, Giovanni Malloy, Lisa Pelled Colabella, Erin N. Leidy, Adaeze Ibeanu, and Madison Williams, *Aging Systems in the Information Age: An Assessment of Technical Debt in Army Enterprise Information Technology*, RAND Corporation, 2024, Not available to the general public.

Woods, David L., John M. Wyma, E. William Yund, Timothy J. Herron, and Bruce Reed. "Factors Influencing the Latency of Simple Reaction Time," *Frontiers in Human Neuroscience*, Vol. 9, 2015.

Whaley, Alison, "Ombuds," DoD Office of Inspector General, webpage, undated. As of December 15, 2025: https://www.dodig.mil/Offices/Ombuds/

Wright, Thomas A. and Russell Cropanzano, "Emotional Exhaustion as a Predictor of Job Performance and Voluntary Turnover," *Journal of Applied Psychology*, Vol. 83, No. 3, 1998.