2024-12

FROM POLICY TO PRACTICE: FACTORS DRIVING CYBERSECURITY IMPLEMENTATION FAILURES

Landis, Christopher B.

Monterey, CA; Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

FROM POLICY TO PRACTICE: FACTORS DRIVING CYBERSECURITY IMPLEMENTATION FAILURES

by

Christopher B. Landis

December 2024

Dissertation Supervisors: Joshua A. Kroll
Britta Hale

THIS PAGE INTENTIONALLY LEFT BLANK

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

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2024	**3. REPORT TYPE AND DATES COVERED** Dissertation

4. TITLE AND SUBTITLE FROM POLICY TO PRACTICE: FACTORS DRIVING CYBERSECURITY IMPLEMENTATION FAILURES	**5. FUNDING NUMBERS** N6600123WX00492
6. AUTHOR(S) Christopher B. Landis	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) NIWC Pacific, San Diego, CA 92152	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This report has supplemental(s).

12a. DISTRIBUTION / AVAILABILITY STATEMENT Distribution Statement A. Approved for public release: Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Organizations develop and implement policies to control their computing systems, turning high-level legal and business requirements into data and software that govern how these systems behave. However, implemented policies may not capture the substance of their high-level natural language source, resulting in a policy-implementation gap. We apply the ambiguity-conflict model of policy implementation to study this gap for cybersecurity in three important cases. First, we examine the nonuse of system resilience capabilities in U.S. Navy tactical networks. Second, we analyze the degree to which standards and frameworks could have helped organizations identify and control risk to support policymakers' intent. Third, we investigate non-conformance with a legally directed cybersecurity policy among U.S. Government agencies. We devised and used a gap measurement model in our analyses. Across these diverse cases that deepen our understanding of the policy-implementation gap in cybersecurity, we find that factors beyond technical capabilities drive gaps more than technology itself. As the separation (especially in organizational structure) between cybersecurity policymakers and practitioners implementing policy increases, so does the opportunity for having a larger policy-implementation gap. A decoupling of computing systems' policy requirements from their implementations, intentional policy ambiguity, and organizational complexity contributed to observed policy-implementation gaps.

14. SUBJECT TERMS cybersecurity, policy-implementation gap, ambiguity-conflict model, organizational factors, information technology, IT adoption, privacy	**15. NUMBER OF PAGES** 237
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

FROM POLICY TO PRACTICE: FACTORS DRIVING CYBERSECURITY IMPLEMENTATION FAILURES

Christopher B. Landis
Commander, United States Navy
BS, United States Naval Academy, 2007
MITS, Carnegie Mellon University, 2013

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
December 2024

Approved by:

Joshua A. Kroll
Department of
Computer Science
Dissertation Supervisor

Britta Hale
Department of
Computer Science
Dissertation Co-Supervisor

Neil C. Rowe
Department of
Computer Science
Dissertation Chair

Chad A. Bollmann
Department of Electrical
and Computer Engineering

Michelle Mazurek
University of Maryland

Cynthia E. Irvine
Department of
Computer Science

Approved by:

Gurminder Singh
Chair, Department of Computer Science

James B. Michael
Vice Provost of Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Organizations develop and implement policies to control their computing systems, turning high-level legal and business requirements into data and software that govern how these systems behave. However, implemented policies may not capture the substance of their high-level natural language source, resulting in a policy-implementation gap. We apply the ambiguity-conflict model of policy implementation to study this gap for cybersecurity in three important cases. First, we examine the nonuse of system resilience capabilities in U.S. Navy tactical networks. Second, we analyze the degree to which standards and frameworks could have helped organizations identify and control risk to support policymakers' intent. Third, we investigate non-conformance with a legally directed cybersecurity policy among U.S. Government agencies. We devised and used a gap measurement model in our analyses. Across these diverse cases that deepen our understanding of the policy-implementation gap in cybersecurity, we find that factors beyond technical capabilities drive gaps more than technology itself. As the separation (especially in organizational structure) between cybersecurity policymakers and practitioners implementing policy increases, so does the opportunity for having a larger policy-implementation gap. A decoupling of computing systems' policy requirements from their implementations, intentional policy ambiguity, and organizational complexity contributed to observed policy-implementation gaps.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

3DES	Triple Data Encryption Standard
ADNS	Automated Digital Network System
AEHF	Advanced Extremely High Frequency
AI	artificial intelligence
AOR	area of responsibility
BOD	Binding Operational Directive
C2	command and control
C4	command, control, communications, and computers
CaaS	Communications as a Service
CANES	Consolidated Afloat Network and Enterprise Services
CBSP	Commercial Broadband Satellite Program
CCPA	California Consumer Privacy Act
CD2E	communications-degraded or denied environment
CENTRIXS	Combined Enterprise Regional Information Exchange System
CEO	Chief Executive Officer
CHOP	change of operational control
CISA	Cybersecurity and Infrastructure Security Agency
CSIO	Combat Systems Information Officer
CTI	controlled technical information
CUI	controlled unclassified information
CVN	aircraft carrier, nuclear powered
DHS	Department of Homeland Security
DKIM	Domain Keys Identified Mail

DMARC	Domain-based Message Authentication, Reporting, and Conformance
DNS	Domain Name System
DoD	Department of Defense
DoS	denial-of-service
DSCS	Defense Satellite Communication System
DTG	date-time group
ECRNOC	European-Central Region Network Operations Center
EHF	extremely high frequency
ERP	enterprise resource planning
FERPA	Family Educational Rights and Privacy Act
GDPR	General Data Protection Regulation
HSTS	Hypertext Transfer Protocol (HTTP) Strict Transport Security
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
IC	Intelligence Community
IP	Internet Protocol
IRB	Institutional Review Board
IT	information technology
JFTOC	Joint Fleet Telecommunications Watch Officer
Kbps	kilobits per second
MitM	man-in-the-middle
ML	machine learning
MOC	maritime operations center
MX	mail exchange
NAVWAR	Naval Information Warfare Systems Command
NCCIC	National Cybersecurity and Communications Integrations Center

NIPRNet	Non-secure Internet Protocol Router Network
NIST	National Institute of Standards and Technology
NIWC	Naval Information Warfare Center
NOC	network operations center
NPS	Naval Postgraduate School
NSS	National Security System
NYC	New York City
OIG	office of the inspector general
OPNAV	Office of the Chief of Naval Operations
OPSEC	operational security
PET	privacy-enhancing technology
PII	personally identifiable information
POR	program of record
PRNOC	Pacific Region Network Operations Center
PSO	predicate singling-out
QI	quasi-identifier
RC4	Rivest Cipher 4
RFC	Request for Comments
RFI	request for information
RSSC	Regional Satellite communications (SATCOM) Support Center
RUA	reporting uniform resource identifier (URI) for aggregate data
RUF	reporting uniform resource identifier (URI) for failure data
SAA	satellite access authorization
SAR	satellite access request
SATCOM	satellite communications
SIPRNet	Secure Internet Protocol Router Network

SP	Special Publication
SPF	Sender Policy Framework
SSL	Secure Sockets Layer
STARTTLS	[a command to] start Transport Layer Security
STEP	Standardized Tactical Entry Point
TDMA	time-division multiple access
TIP	Time-division multiple access (TDMA) Interface Processor
TLC	Taxi and Limousine Commission
TLD	top-level domain
TLS	Transport Layer Security
TYCOM	type commander
URI	uniform resource identifier
USG	U.S. Government
WAN	wide-area network
WGS	Wideband Global System

Acknowledgments

Soli Deo Gloria

> Many scientists these days extol the virtue of basic research which has no apparent purpose other than to reveal some aspect of nature. In eras of cost-cutting, they urge the funders to support them because who knows what fruits basic research may eventually produce. . . Why not say, instead, as our forefathers did, that basic research is one of the finest ways we have of glorifying God?
>
> —C. West Churchman [1, p. 244]

Earning a Doctor of Philosophy has been the greatest challenge I have faced. Some may challenge this claim by asking about my basic military training or experiences while deployed in harm's way. I concede that those and some other life experiences have included some very intense, trying times; however, the years-long duration causes the PhD process to surpass those other, more intense but shorter trials. This has been a rollercoaster of achievement and debilitatingly stressful doubt. I have many to thank for their support at various points throughout this journey. "The horse is prepared for the day of battle, but victory belongs to the LORD" (Proverbs 21:31, NASB).

Committee

Thank you to Dr. Joshua Kroll for serving as my advisor throughout this adventure. There have been many times of doubt and struggle unrelated to the content of this dissertation through which you helped me navigate. Thank you for helping me chart my own course and for devoting more time to your advisees than I hear many others receive from their advisors. I have learned much from you. You have a brilliant mind for bringing together seemingly unrelated concepts to form influential interdisciplinary work.

Thank you to Dr. Britta Hale for serving as my advisor when Dr. Kroll was on parental leave. Moreover, thank you for your constructive feedback on my dissertation that helped me improve the work, especially regarding the interpretation of technology-use mediation for Chapter 3 and the assumptions of statistical tests for Chapter 5.

Thank you to Dr. Cynthia Irvine for sharing some of your wisdom with me that you have acquired across your decades of professional academic tenure. If circumstances would have allowed, I would have liked to have spent much more time with you, to gain more from your insight and wisdom.

Thank you to Dr. Chad Bollmann for bringing your perspective as a Permanent Military Professor (PMP) and for helping me to maintain a greater perspective during the tougher times. Also, thank you for identifying for me a contribution that I had not even realized I had made and helping me to develop it. Seeing this enabled me to expand on the concept and improve the value of this work.

Thank you to Dr. Michelle Mazurek for providing a non-Naval Postgraduate School (NPS) perspective, participating in my committee from the University of Maryland. Ever since I first heard your keynote address [2], I knew I wanted to work more in your line of research. I look forward to working with you from the U.S. Naval Academy (USNA).

Thank you to Dr. Neil Rowe for agreeing to join my committee later in this process. I appreciate your support in this process, your perspective on my experiences and work, and your sense of duty to the NPS mission. The quality of my writing improved greatly under your leadership.

Other Professionals

Thank you to Dr. Gurminder Singh, Dr. Sandi Leavitt, and Dr. Thor Martinsen for listening, providing advice, and supporting me throughout this journey.

I am grateful for the NPS Graduate Writing Center workshops and writing coaches that have helped me improve my writing, in particular, Dr. Kate Egerton in 2022 and Ms. Alison Scharmota in 2024. Also, I found Creswell and Báez's practical book on qualitative research very helpful in developing my writing [3].

I am grateful for the advice and insight on statistics from Dr. Tom Lucas and Dr. Robert Koyak.

Thank you to Ms. Janice Long, NPS Thesis Processing Office, for letting me volunteer to fix, improve, and maintain this template. Working on the template for "a little here and a

little there" over this long span of time has put my fingerprints on much more of the template than I would have anticipated from the start. I have enjoyed using my talents to serve others in this way. Your clear vision for the template was invaluable in guiding its development.

Thank you to my fellow PhD students for your support, encouragement, and recommendations, especially Dr. Aaron Geary, Dr. Dan Lukaszewski, Dr. Ed Jatho, Dr. Pedro Ortiz, and Dr. Dominiqué Hittner who were ahead of me, CDR Jon Durham and Mr. Dave Justamante who will graduate with me, and CDR Peter Barkley, MAJ Tim Walsh, Mr. Jason Landsborough, and CDR Rich Thompson who will soon graduate.

Chapter Acknowledgments

For Chapter 3, "Adopting Resilient Information Technology Afloat" [4]: I thank Ms. Carly Jackson for her official endorsement of my research surveys to meet the Navy Survey Office's eligibility requirements and the groups of 69 and 4 Information Professional Officers that respectively participated in my two surveys, shown in Appendices A and B. For their specific technical assistance from personnel at Naval Information Warfare Center (NIWC) Pacific, I especially thank Mr. Pros Loung and his Automated Digital Network System (ADNS) team, Mr. Chinh Nguyen, and Mr. Phil Parlin. I thank the anonymous reviewers for their constructive feedback. I also specifically thank Dr. Emanuel Moss, Dr. Rene Rendon, Dr. Jamie Porchia, and Mr. Howard Pace for their feedback on a draft of this work.

For Chapter 4, "Using a Commonly Accepted Framework to Protect Against Privacy Inferences" [5]: I thank the anonymous reviewers for their constructive feedback. I appreciate the time and feedback that the privacy engineering group at the National Institute of Standards and Technology (NIST) gave to us in discussing this work, especially Ms. Naomi Lefkovitz and Mr. Dylan Gilbert. I also specifically thank Dr. Nitin Kohli and Dr. Chris Hoofnagle for their feedback.

For Chapter 5, "Implementing an Enterprise Cybersecurity Policy across the U.S. Federal Executive Branch:" Thank you to LT Thomas Glade, who provided the Web server probing raw data for me to incorporate into my analysis [6]. I thank the ten U.S. Government (USG) agency representatives who participated in our survey, shown in Appendix C, and two Cybersecurity and Infrastructure Security Agency (CISA) employees who agreed to discuss the work. I thank the anonymous reviewers for their constructive feedback.

Thank you to the NPS Institutional Review Board (IRB) for their time and assistance in helping Dr. Kroll and me ensure all the surveys mentioned here complied with all applicable human subjects research (HSR) requirements.

NIWC Pacific

NIWC Pacific funded a portion of this research as part of the NPS-NIWC Pacific Research Fellowship, which I conducted with the assistance and under the mentorship of Dr. Stephen Dabideen, NIWC Pacific. Thank you to Dr. Dabideen for your guidance that, in part, has led to the work I present here in Chapter 3 and in other venues [4], [7]. While at USNA, I would like to continue research that improves aspects of the Navy's afloat Internet Protocol (IP) tactical networks.

Conference Attendance

Thank you to Dr. Hale for funding my travel to and attendance at the Association for Computing Machinery (ACM) Conference on Computer and Communications Security (CCS) in November 2022 in Los Angeles, CA.

Thank you to Dr. Irvine for funding my travel to and attendance at the Institute of Electrical and Electronics Engineers (IEEE) Conference on Security and Privacy (S&P) in May 2023 in San Francisco, CA.

Thank you to NIWC Pacific for funding my travel to and attendance at the IEEE Conference on Military Communications (MILCOM) in October–November 2023 in Boston, MA, at which I presented an in-progress status of my work [7].

Thank you to the NPS Department of Defense Management's Acquisition Research Program that sponsors NPS student attendance at the Annual Acquisition Research Symposium in Monterey, CA, in whose May 2024 proceedings a portion of my work appears [4].

Thank you to Dr. Kroll for funding my travel to and attendance at the Privacy Enhancing Technologies Symposium (PETS) in July 2024 in Bristol, United Kingdom, at which I presented a portion of my work [5]. I am also grateful for the opportunity that you provided to me during this trip to present my dissertation work to faculty at University College

London, including Dr. Michael Veale (as the host), Dr. Reuben Binns, Dr. Tristan Caulfield, and Dr. Jat Singh.

Thank you to Dr. Kroll for funding my travel to and attendance at the USENIX Symposium on Usable Privacy and Security (SOUPS) and USENIX Security Symposium in August 2024 in Philadelphia, PA.

Personal

To my beloved bride and best friend, Amy: It seems I cannot say anything but can praise and thank God for bringing us together. Neither of us could have planned or predicted how the challenges of our three years in Monterey would draw us closer to each other and to God. I love you and look forward to our next adventure together.

To my children: Thank you for bearing with me through these years when I have had to tell you repeatedly, "No, I can't play with you or read to you now because I need to get this work done." Now is the time for me to catch up on these important aspects of being your father.

To my parents: How could this work (and my life) have played out this way had it not been for your life-long love, demonstration and appreciation of hard work, and supporting encouragement? Thank you. I love you, too.

Thank you to the many other family members and friends who have been praying for and supporting me through this process.

> I will lift up my eyes to the mountains; from where shall my help come? My help comes from the LORD, who made heaven and earth. (Psalm 121:1–2, NASB)

List of References

[1] C. W. Churchman, *The Design of Inquiring Systems: Basic Concepts of Systems and Organization*. New York: Basic Books, 1971.

[2] M. Mazurek, "We are the experts, and we are the problem: The security advice fiasco," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and*

Communications Security, 2022, p. 7. Available: https://doi.org/10.1145/3548606.3559394

[3] J. W. Creswell and J. C. Báez, *30 Essential Skills for the Qualitative Researcher*, 2nd ed. Los Angeles, CA: Sage Publications, Inc., 2021.

[4] C. B. Landis and J. A. Kroll, "Acquired and deployed but not adopted: Lost mission effectiveness without resilient chat afloat," in *Proceedings of the Twenty-first Annual Acquisition Research Symposium*, 2024, vol. 3, pp. 199–214. Available: https://dair.nps.edu/handle/123456789/5161

[5] C. B. Landis and J. A. Kroll, "Mitigating inference risks with the NIST privacy framework," *Proceedings on Privacy Enhancing Technologies*, vol. 2024, no. 1, pp. 217–231, 2024. Available: https://doi.org/10.56553/popets-2024-0013

[6] T. L. Glade, "Comparing TLS and certificate configurations of government and non-government websites," M.S. thesis, Naval Postgraduate School, Monterey, CA, June 2024. Available: https://hdl.handle.net/10945/73127

[7] C. B. Landis, "Afloat resilience of IP tactical communications: Chat systems and sailors," Lightning talk presented at IEEE Conference on Military Communications (MILCOM), Boston, MA, October 31, 2023. DTIC citation AD1215296, restricted distribution. Available: https://milcom2023.milcom.org/program/restricted-technical-program

CHAPTER 1:
Introduction

In April 2021, the Colonial Pipeline company's computer network suffered a ransomware attack, leading the company to shut down the pipeline, depriving the eastern U.S. of more than 100 million gallons of fuel oil daily and impacting the economy [1]. The Colonial Pipeline ransomware attacks relied on a well understood cybersecurity weakness for gaining initial access: a stale account configured for single-factor, password authentication. Why do organizations continue to struggle to overcome longstanding weaknesses like these? Because these weaknesses are well understood, plenty of cybersecurity tools (for example, online technical repositories, professional certification guides, and textbooks) can help system administrators with cybersecurity management and commonly accepted standards and frameworks inform readers of best practices. Were these tools or standards used? If so, did they help organizations establish and implement cybersecurity policies to mitigate these and other vulnerabilities?

Organizations develop and implement *policies* to affect the operation and configuration of these systems to protect their information and achieve their goals, such as increased mission effectiveness or assurance. Organizations sometimes experience unexpected and undesirable outcomes to their cybersecurity policy decisions related to the operation and configuration of their software systems. These could be for large-scale breaches of sensitive data or consumer accounts [2], [3] or fake emails from a trusted domain or business email compromises [4], [5]. Other examples include gaps between public claims and actual data-sharing activities [6], [7] and cases where inferences from non-sensitive data about consumers reveal facts those consumers would likely not consent to disclose [8].

People intend for security policies to create positive security outcomes, but gaps persist when translating organizational and policy ambiguities into concrete implementations. Establishing the goals for security policy at an organization's highest level differs from implementing security policy at lower levels. Because goals are often expressed in policies (or implicitly as an organization's collective actions, which we treat as an implicit form of policy), we borrow ideas from the study of policy-implementation failure for describing the loci

and etiologies of cybersecurity failure. Specifically, we apply Matland's *ambiguity-conflict model of policy implementation*, which reconciles literatures on top-down and bottom-up policy implementation approaches to identify the necessary conditions for successful policy implementation [9]. By examining diverse cases, we identify the key leverage points for improving implementation while creating space for discussions of policy-implementation effectiveness. Besides bringing a policy's realized implementation closer to the intended implementation, benefits of closing the gap include better consistency with the goals from which the organization derives its policy and greater uniformity in implementation across an enterprise.

If a cybersecurity-policy implementation outcome differs from that intended such that an organization does not achieve its goals, the organization can suffer harms in the gap between its intended and realized policy implementations, as in the examples we will discuss. Such a gap between the intended and realized outcomes can occur regardless of the policy's ambiguity and can occur by over-fitting or under-fitting an implementation to a policy. It can also occur even when organizations execute each step of their policy implementation according to the best standard practices, which many consider to be proper translations of requirements.

We call this gap between the effects intended and those realized a *policy-implementation gap* [10]. A gap in cybersecurity impedes an organization in achieving its goals and leaves it vulnerable to compromise. By *organization*, we mean "a system of consciously coordinated activities or forces of two or more persons" [11, p. 73]. Further, it should have "a defined mission/goal and a defined boundary, using information systems to execute that mission, and with responsibility for managing its own risks and performance" [12, p. B-4]. This definition labels both the U.S. Navy and the Department of Defense (DoD) as organizations, even though the Navy is part of the DoD. We anthropomorphize organizations with human attributes (like their goals) as shorthand for referring to, for example, the goals agreed upon by the people comprising the organizations. Organizations can also have emergent behavior not present in their parts.

A high-level goal of organizations is to decrease the frequency and severity of policy-implementation gaps in organizations of all types and sizes to avoid incidents, including data breaches [2] and lost revenue due to misconfigurations [5]. Taking a step toward this

goal, we aim to identify factors contributing to policy-implementation gaps in cybersecurity. By systematically identifying factors that lead to policy-implementation gaps, we can form interventions for reducing the impact of these factors on the gap. This brings us to our primary research question:

RQ 1. What non-adversarial sources can cause inconsistencies in implementation and the realized outcomes that differ from a cybersecurity policymaker's stated intentions?

Matland synthesized studies on the policy-implementation gap, building the ambiguity-conflict model as a way by which we can interpret how the policy-implementation gap manifests [9]. His analytical work allows him to define conditions under which policy implementation succeeds or fails. We describe this model in greater detail in Chapter 2 and apply it to our studies in Chapters 3, 4, and 5.

We use a simple model of system life cycle as comprised by three phases (adoption, continuance, and discontinuance), following work by Bölen, Calisir, and Özen synthesizing many models used across the systems engineering and information systems literatures [13]. *Adoption* consists of organizations' and users' decision calculi to use the new information system and their initial assessments on how well it serves their purposes, that is, its acquisition, implementation, and initial user training. Identifying a problem, defining requirements, and selecting an information system all factor into the decision calculi. *Continuance* represents organizations' and users' information system use, that is, its operation, sustainment, and maintenance. *Discontinuance* starts when, for any reason, organizations and users stop using the information system, which typically involves transitioning to the adoption phase of the information system's replacement [13].

Using a vignette of common organizational actions in this life cycle to guide our inquiry (expanded in Section 1.2), we decomposed the main research question into several, more specific research questions within the life cycle's adoption and continuance phases, as shown in Figure 1.1. At a high level, the vignette describes an organization that identifies a problem, decides to solve that problem by adopting an information technology (IT) system, develops requirements, and translates them to practice in the adoption phase. Here, the requirements may define or even be expressed as policy. In the continuance phase, the

organization operates and maintains the system and verifies that it is meeting the operational requirements. Later, the organization may discontinue the system.

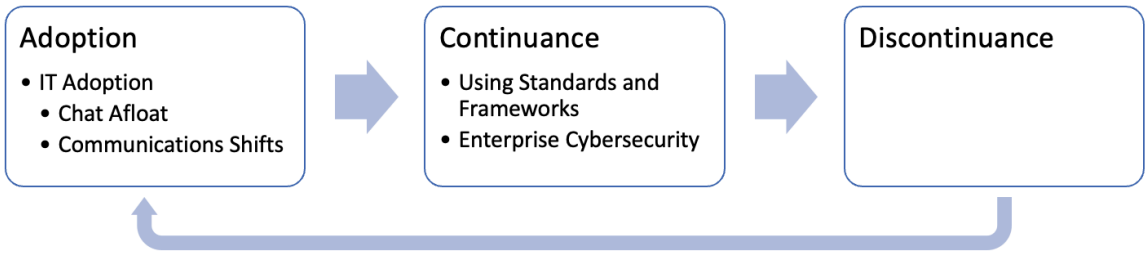


Figure 1.1. The Information Systems Life Cycle and Our Selected Studies. Adapted from [13, p. 551].

Organizations commonly develop or adopt information systems to achieve their goals, establishing policy to affect their operational procedures and technical configurations. We focus on the adoption of IT capabilities aboard U.S. Navy ships, specifically related to sailors' use of resilience-providing capabilities in afloat tactical networks for chat systems and communications shifts. The Navy paid for these capabilities, intending (by policy) that the capabilities would improve warfighting effectiveness, but has not implemented them to that effect.

RQ 2. How can the level of adoption of available IT capabilities contribute to gaps between policy and practice?

To make good risk decisions, organizations must identify risks, but can face challenges doing so comprehensively. Organizations customarily use commonly accepted standards and frameworks for improving the operation and configuration of their software systems. These sometimes take the form of a checklist for compliance and assessment. Risk-assessment standards are a focal point for organizing and documenting the work of risk perception in two senses: translating risk reduction goals to requirements and controls (using checklists as risk perception tools) and measuring how much organizations' policy implementations meet their requirements (using checklists as minimum performance standards). Under both kinds of assessment (checklists as risk perception tools and minimum performance standards), standards and frameworks offering checklists of controls help with translating and assessing

requirements at the lower levels of implementation. Complicating their use, however, some requirements from these standards and frameworks are untestable at the component level but emerge at the system-level, like security or privacy. Checklists, therefore, can falsely imply the achievement of emergent properties.

We focus on organizations' use of standards and frameworks for privacy protections because of the privacy impact in the cited incidents involving compromising email accounts [3], [5] and inferring sensitive attributes about people [6], [7], [8]. Privacy is more than data protection, though. Data protection typically relies on confidentiality and information-flow measures like access controls. While data protection is an important part of privacy, privacy also concerns attributes and identities that one can learn from data to which one lacks access. Access-control and information-flow approaches are marginally related to inferences and do not completely address them. Standards and frameworks can help draw an organization's attention to these types of privacy risk.

RQ 3. How can standards or frameworks guide or hinder organizations' policy implementation?

Organizations select, deploy, and implement cybersecurity policies to affect associated technical controls and implementations. Commonly, organizational hierarchies have sub-organizations within organizations, sometimes with formal boundaries between layers to enable a subordinate's autonomy. Taking the U.S. Navy and DoD example a step further, consider the Executive Branch of the U.S. Government (USG), of which the DoD is a member. The Executive Branch is an organization whose components are relatively independent of each other. An Executive Branch policy can require implementation by all its components, yet the organizations that together comprise the Executive Branch are individually responsible for implementation. These Executive Branch agencies, in turn, act as their own organizations, directing their components to implement these directed policies. In this context, we ask:

RQ 4. What can cause the implementations of a mandatory cybersecurity policy to vary across U.S. Federal Executive Branch organizations?

Because these questions involve assessing a degree of policy implementation, we consider it worthwhile to develop a model for conceptualizing such a measurement.

RQ 5. How can we assess the magnitude of a policy-implementation gap?

To answer these questions, we conduct a series of analytical and empirical studies exploring the factors contributing to the policy-implementation gap across various design issues and system-application types. First, we investigate how administrators use of certain U.S. Navy communications systems without resilience-providing configurations serves the stated high-level goal of improving communications resilience. Second, we investigate the relationship between standards such as checklists and the ability of organizations to gather information about their level of privacy risk, where privacy is a goal not reducible to a property of any specific system component, while checklists focus on component-level behavior. Finally, we investigate observationally the variation in the policy-implementation gap in a situation where the same policy applies across many independently governed organizations, and the outcome of policy implementation is readily measurable at the technical level. These studies give us a descriptive view into the nature of variability of the policy-implementation gap, grounding subsequent work to identify and isolate causal factors and interventions to minimize it. We thus offer the following **primary contributions**:

1. Using the ambiguity-conflict model for analyzing the policy-implementation gap and, for the first time, for organizational cybersecurity policy, our diverse studies broaden the model's usefulness and offer an alternative perspective for understanding the policy-implementation gap in cybersecurity.
2. We conducted two studies on IT capability nonuse not yet addressed by the Navy by surveying administrators and integrating technology-use mediation concepts with the ambiguity-conflict model.
3. We examined the conditions for privacy-risk perception and mitigation in real privacy incidents against a standard privacy framework, publishing an analysis of how the National Institute of Standards and Technology (NIST) Privacy Framework could have more adequately helped to identify the risks exploited in the studied incidents.
4. By gathering data on both network-services and information-security workers, we explored how the policy-implementation gap manifests across the U.S. Federal Executive Branch when implementing a legally mandated cybersecurity policy.
5. We devised a model for organizations to measure the policy-implementation gap and demonstrated how to use the model for measuring the gap across an enterprise cybersecurity policy's implementations in the USG.

This dissertation contributes to the field of organizational cybersecurity, whose members study the "efforts organizations take to protect and defend their information assets, regardless of the form in which those assets exist, from threats internal and external to the organization" [14, p. 4]. The field integrates knowledge from computer and social sciences and offers a more holistic perspective of the policy-implementation gap than is usually considered [10]. Our work contributes to the study of organizational cybersecurity by exploring shortfalls in cybersecurity policy implementation in organizations.

We begin our investigation of the policy-implementation gap by first defining *organizational cybersecurity policy*. We then explore the policy-implementation gap within established life cycles on how organizations adopt, operate, and configure information systems to achieve their goals.

1.1 Organizational Cybersecurity Policies

The cybersecurity policies that we study in this dissertation are organizational policies and their implementations, which comprise the lower levels of a policy hierarchy, shown in Figure 1.2. A *policy* is "a course or principle of action adopted or proposed by a government, party, business, or individual" [15]. Organizations use policies to bridge goals with implementable technical-protection needs. People's and organizations' *goals* influence policy at each level, and policies influence people's goals by providing requirements, means, and constraints [16]. For example, the goals of those making public-policy influence future public policies, which may in turn influence future policymakers' goals. Governments enact *public policies* in the form of laws and regulations. In turn, organizations form their own policies that balance organizational goals with complying with public policy. *Organizational policies* describe an organization's specific approach to a principle or function, as approved by the organization's management. For example, an organization's email policy may direct how the organization's members are to use email, may describe prohibited activities, and may specify actions to ensure the email system's security as part of the organization's information system. The aspects of organizational policy can vary from ambiguous principles to objectively verifiable actions. When policies are ambiguous, they require the sound judgment of the organization's practitioners for their implementation and are otherwise incommensurable [17].

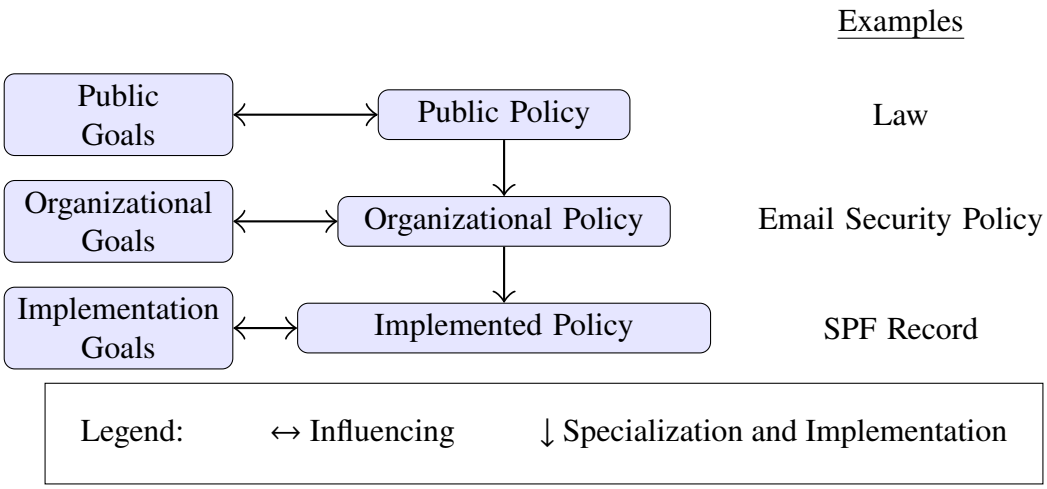


Figure 1.2. A Policy Hierarchy Model. Lower-level policies specialize and implement higher-level policies. At all levels, goals influence policymakers and practitioners implementing policy.

Implemented policies further specialize and implement organizational policies, refining organizational goals while balancing the goals of the practitioners who implement the policy [9]. Residing at the lowest level of our policy hierarchy, implemented policies refer to fully specified technical requirements, like a firewall access control list (ACL) or a Domain Name System (DNS) text record representing a Sender Policy Framework (SPF) policy. Implemented policies in an information system's software components both afford and constrain the implementation of future policies. As such, the *implementation goals* consist of trade-offs between the constraints and capabilities of the implemented system. Cybersecurity practitioners work within these affordances, constraints, and the competing demands on their time and talent to implement the organization's policies by technical means.

While some cybersecurity policies direct precise requirements, others have ambiguous requirements. When the requirements are precise, like "use Transport Layer Security (TLS) version 1.2 or 1.3" (because of [18], for example) evaluation becomes easier, especially with measuring the policy's implementation by automata. The U.S. Department of Homeland Security (DHS) Binding Operational Directive (BOD) 18-01 (the subject of Chapter 5) is a cybersecurity policy with precise requirements [19]. To evaluate agencies' conformance

to it, we use the model in this section to devise a *distance to conformance* model that indicates the number of nonconforming policy requirements. On the other hand, cybersecurity practitioners must apply their professional judgment in implementing ambiguous policy requirements, like "use secure communications protocols." This is because many such protocols exist, and professionals must operationalize guidance according to specifics of an application. For example, many of the U.S. DoD Operations Security (OPSEC) Program policy's requirements (mentioned in Chapter 4) are ambiguous and need practitioners' professional judgment for their intended implementations [20]. Such individual interpretations can lead to differing implementations of the same policy.

One might think that increasing policy requirements' precision would address this problem. Requirements engineers have tried this through formal modeling methods [21], [22]; however, policies are often purposefully ambiguous to enable flexibility or make space for practitioners' professional judgments during implementation. Separate from implementation, quantitatively evaluating ambiguous policy requirements is challenging because of a natural lack of consensus of their meanings. Encoding them into predicates and determining how to evaluate the predicates are challenging tasks. Moreover, resolving ambiguity into something that can be evaluated may not capture the intended sense of the ambiguous language. Practitioners must also use their professional judgment when completing implied tasks that the policy does not specify, like acquiring a certificate for TLS from a reputable certificate authority. These implied tasks are common in policies, independent of their requirements' precision.

Practical cybersecurity work revolves around the implementation and enforcement of policies. Policies define (1) the goals for system operation and protection, (2) the nature of the trust relationship between the system and the people and organizations that depend on it, and (3) the threat model in which systems live [23]. Sometimes organizational cybersecurity policies specify tools or procedures to be implemented. Implemented policies vary by representation, from code that performs specialized calculations to data. *Security policy* has a precise etymology in computer-research literature [23]; organizations operationalize their policy objectives by establishing security policies and by building technological artifacts and processes that implement (and sometimes automate the enforcement of) these policies. Although prior work has translated implemented policy into computer-interpretable/executable formats like context-free grammars [24], [25], relatively few have

made progress in translating the natural language of organizational policy for specific use cases into computer-verifiable implemented policy [26].

If organizations could implement outcomes, they choose a policy of no cybersecurity incidents. Instead, policy implementations influence outcomes. That is, the gap is between the policy and the implementation; the outcomes happen separately, but indicate a policy-implementation gap when the organization does not achieve its cybersecurity goals, as shown in Figure 1.3. To understand the intermediate steps in the policy-implementation gap, we need an ontology to translate semantically between a policy's requirements and the controls selected to achieve those requirements. Let us imagine that such a semantic ontology exists for validating cybersecurity controls. How can we prove that controls are necessary and sufficient? Proving that a cybersecurity control is insufficient is possible, and sometimes trivial. Proving the opposite, that a cybersecurity control is necessary and sufficient to improve or achieve a secure outcome idealized in a policy, is unfalsifiable [27]: Any relevant incident disproves a claim that a policy "secures" a system, and one cannot prove such a security claim because of unknown vulnerabilities [28].

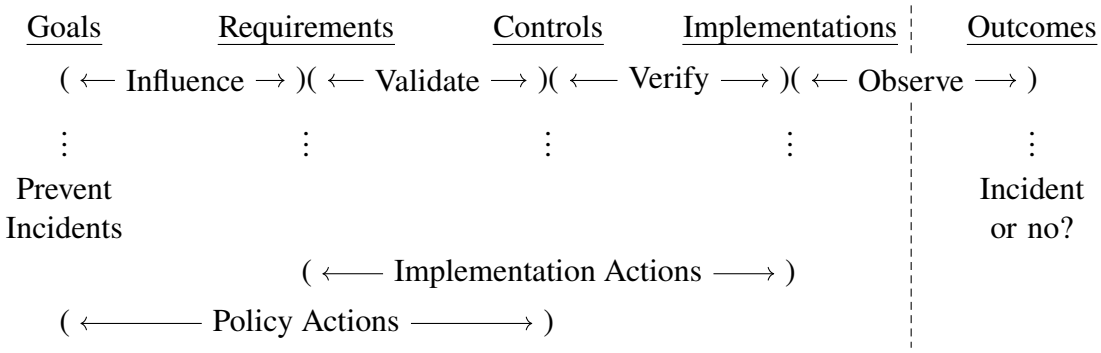


Figure 1.3. Relations in Translating Policy. The policy-implementation gap exists between policy actions and implementation actions but is detected across the span between goals and outcomes. Although an organization's policy implementation affects outcomes, the organization does not control them directly as it does the other columns in this model.

In the model that we use to define the terms and their relationships (shown in Figure 1.3), goals represent the organization's cybersecurity policy objectives, requirements must fulfill

the organization's goals for that policy, and outcomes must be relevant to that policy. For example, an organization's goals for an email security policy could include, "Ensure the authenticity of users' email messages." We assume that the requirements are consistent with achieving the policy's goals and that the scope of the considered outcomes is relevant to the organization's goals for the policy. Once practitioners have implemented the controls, one can verify that their implementation is correct according to each control's specification. Observing a policy's outcomes depends on one's ability to detect a policy-implementation failure (for example, by experiencing an incident), which can indicate that a policy-implementation gap exists. Without such an indicator, we may be unable to recognize that a policy has not achieved its goals.

Organizations commonly separate the work of policy determination from implementation. As a result, *determining* a policy can be outside the remit of technical work. This causes policymakers to lose visibility of the consequences of adopting a particular policy [29], and it engenders ambiguity-induced and conflict-induced difficulties for implementers, which Matland's model addresses [9]. Such a situation incurs more risk than intended by the organization's policymakers. By *risk*, we mean *the potential for incurring adverse effects on organizational operations, assets, and individuals, instigated by threats to information systems*, which is consistent with the term's use throughout the NIST Special Publication (SP) 800-39 [12] and is appropriate for our studies involving government entities.

Although both academic research and practical cybersecurity work assume that organizations choose security policies to improve security outcomes, various human, organizational, and other factors influence these policies, sometimes in ways inconsistent with security goals. *Human factors* can include attitudes, level of knowledge, incentives, work practices, and error [30], [31], [32]. *Organizational factors* can include structure, incentives, work practices, and risk perception [29]. Technical policy *implementation* is generally consigned to specialized cybersecurity practitioners who have the technical knowledge to interpret and implement policies but not the organizational power to balance cybersecurity needs with other organizational needs [33].

We argue that this dichotomy between determining and implementing policy aggravates the policy-implementation gap. It can foster a bivalent dynamic with poor security outcomes: Managers declare policies that do not in fact result in their intended security outcomes,

while the complexity of integrating technology into organizations hinders cybersecurity practitioners' efforts to bridge this gap [34]. A breakdown in or lack of communication between policymakers and practitioners can have a similar effect. The result is policies that poorly fit an organization's threat environment or that encourage practices known to be bad for the organization's security or the people inside it.

1.1.1 Measuring Policy Conformance

One can measure a policy-implementation gap with a measurement model. One can model a list of 1 to M policy requirements as a checklist (which we discuss further in Sections 2.3 and 4.4.4) that forms a data structure of conjoined predicates and evaluates as true only if all requirements are satisfied:

$$Req1 \;\&\&\; Req2 \;\&\&\; \ldots \;\&\&\; ReqM == \text{TRUE}|\text{FALSE}. \tag{1.1}$$

Thus, FALSE indicates that a gap exists; however, TRUE cannot always indicate the absence of a gap just as we cannot prove a cybersecurity policy is secure [27]. This is problematic for those who believe completing a cybersecurity checklist will make their systems secure.

Policies typically aim to achieve system-level outcomes (e.g., no cybersecurity incidents) through component-level configuration changes. As such, organizations implement capabilities at the component level to promote the property's emergence at the system level. In practice, organizations contextualize cybersecurity within a threat model that describes specific adversary capabilities against specific cybersecurity goals. Organizations can test whether a cybersecurity policy satisfies the conditions of their model. A greater challenge is validating that the model represents reality, that it truly captures adversaries' capabilities that matter to the organization. Because organizations lack complete knowledge of their adversaries' capabilities, organizations must make assumptions about their adversaries on which their model depends. Policy conformance is a direct representation of system state versus articulated goals but only a proxy for cybersecurity incident risk; validating that a threat model captures meaningful risk avoidance is epistemically problematic [35]. In other words, we can test the precise requirements of a cybersecurity policy (like the TLS example in the opening paragraph of Section 1.1) against an organization's threat model but not that the requirements make a cyber-system secure. Cybersecurity is therefore an

emergent property, that is, it occurs only as a product of the interactions of systems, societies, and their components. Other emergent properties, sometimes called *non-functional* or *non-functionally-testable* properties, include resilience, privacy, and safety. Our claims on achieving these properties are limited by the models that organizations use in creating their policies.

If we average the Boolean policy conformance statuses of the policy requirements in Equation 1.1 ($\{Req1, Req2, \ldots, ReqM\}$) instead of conjoining them, we form Equation 1.2. With Equation 1.2, we calculate a 0-to-1 status to indicate the degree of policy implementation. This provides a more nuanced policy-implementation conformance model than Equation 1.1. Let $cip(I, P)$ represent the degree from 0 to 1 to which implementation I conforms to policy P where 0 replaces FALSE (implementation has no progress toward policy aims) and 1 replaces TRUE (perfect implementation according to the policymakers). Because policies commonly list multiple requirements for practitioners to implement, let P_k represent requirement k (an integer from 1 to M) of policy P such that $P = \cup_{k=1}^{M} P_k$. Let each P_k have an implementation item, I_k, that implements P_k such that $I = \cup_{k=1}^{M} I_k$. The $cip(I_k, P_k)$ notation (each having a 0 or 1 value) replaces the Boolean policy conformance status notation of the left side of Equation 1.1. Calculating $cip(I, P)$ in the [0–1] range requires averaging the requirements' conformance:

$$cip(I, P) = \frac{\sum_{k=1}^{M} cip(I_k, P_k)}{M}. \tag{1.2}$$

Organizations can use weights in this model. Further developing Equation 1.2 to include a positive-signed weight, w_k, for each P_k establishes a relative importance among the policy requirements:

$$cip(I, P) = \frac{\sum_{k=1}^{M} w_k \cdot cip(I_k, P_k)}{\sum_{k=1}^{M} w_k}. \tag{1.3}$$

Organizations must determine who will assign the values for the weights and what criteria they will use. An organization's policymakers can assign the values according to their priorities for the organization or practitioners can assign them according to their perception

of how the policy requirements protect against the organization's cybersecurity threats. Alternatively, policymakers and practitioners can collaborate to form consensus. Regardless, the values of the weights are subjective, dependent on the expertise and judgment of the people assigning them. Making the sum of the weights equal 1 ($\sum_{k=1}^{M} w_k = 1$) requires additional planning but simplifies Equation 1.3 by making the denominator equal 1.

To use this model for assessing a policy's implementation, an organization can follow these steps:

1. List a policy's requirements and define the criteria for assessing whether an implementation item conforms to the policy requirement.
2. For each policy requirement, assess whether (or the degree to which) an implementation item conforms to it to determine the value of $cip(I_k, P_k)$.
3. Calculate $cip(I, P)$ using Equation 1.2 or 1.3.
4. If measuring multiple implementations of the same policy across an organization's components, organizations can compare each $cip(I, P)$ measurement and average them to form an aggregate $cip_A(I, P)$.

In Step 1, organizations may find that their policy requirements and implementation items do not have a one-to-one relation. We must consider two aspects of this relation. One, although we defined I_k as an implementation item that implements P_k, I_k may not be the only implementation item that could implement P_k. For example, two different implemented configurations may conform to the same policy requirement. Such a situation does not invalidate the model. Two, some implementation items or some policy requirements may overlap. Such overlap may provide indications of the implementation items or policy requirements that are more important to achieving the organization's goals.

If multiple implementation items are necessary for implementing a policy requirement, organizations have at least two options: They can subdivide the policy requirement into multiple requirements to match the separate implementation items or they can consider each implementation item as a step in the implementation process, which we address in the next paragraph. From the other perspective, an implementation item may contribute to multiple policy requirements simultaneously. In such a case, I_k and I_{k+1} may both refer to the same implementation item. This does not necessarily invalidate the model but means that this implementation item will have more weight than other implementation items in

the calculation of $cip(I, P)$. To avoid this imbalance, organizations can combine policy requirements that share an implementation item or adjust the policy requirements' weights accordingly.

As Step 2 offers, the assessment of whether an implementation item conforms to a policy requirement can be a ratio rather than a Boolean value. To use a meaningful ratio, one would need a formal definition of implementation progress, which can be steps in a procedure. For example, if 3 of 4 equally-weighted implementation steps are complete, then $cip(I_k, P_k) = \frac{3}{4}$. When the implementation progress or steps are complete, then $cip(I_k, P_k) = 1$. We could also interpret the ratio as the fraction of the total time necessary to finish implementing the requirement. The formal definition for determining a value to represent an implementation's progress is subjective, dependent on the people defining and applying the measurement model. Such subjectivity can lead to measurement inconsistencies across multiple practitioners and across measuring multiple instances of a policy's implementation, like across a large organization's components. This is one reason why consistency throughout the measurement process, including in the assumptions practitioners make, is important to this model's value.

In Chapters 3 and 5, we use this model for measuring policy conformance. This model helps organizations represent the policy-implementation gap between the policy actions and implementation actions, that is, spanning requirements through implementations, as shown in Figure 1.3. P and I are sets of incompatible types. The sets' incompatible types prevent this model from becoming a metric because a metric requires defining a distance between points of identical types in a metric space [36, p. 1].

Organizations must establish effective cybersecurity and privacy policies that reduce the policy-implementation gap. Our conformance measurement model can help for measuring a policy's implementation progress and comparing implementation progress across implementations of the same policy. When organizations define clear assessment criteria, each $cip(I_j, P_k)$ value becomes less arbitrary, thus giving the model greater meaning to the organization. As organizations establish and implement cybersecurity policies to protect their missions, employees, and customers by better identifying and managing risk, they must continuously reevaluate the validity of their goals and the model of their adversaries' capabilities.

1.2 Exploring the Policy-Implementation Gap

Using a vignette of common organizational actions in the information-systems life cycle to guide our inquiry, we interleave descriptions of our studies and scenarios from the vignette introduced above Section 1.1. Each study offers a different perspective on the policy-implementation gap in cybersecurity and yields its own contributions that contribute to this dissertation's primary contributions.

Computer scientists more commonly reference other life-cycle models. The software-development life cycle is usually modeled as a waterfall or agile process and involving validation and verification processes on software artifacts, as in the V-model [37, pp. 108–116]. Software-development life cycles are important models for actions that occur before the information-systems life cycle, transitioning on their completion into the latter's adoption phase [13]. Similarly, the Control Objectives for Information and Related Technology (COBIT) framework group these together as the "Deliver and Support" and "Monitor and Evaluate" phases [38]. Also related to software development, the specification-implementation gap in formal methods describes "a mismatch between the design and the implementation" [22, p. 340], like the policy-implementation gap in that they both manifest when the intended and realized outcomes do not match. These other models are also better for relating to aspects beyond the scope of this dissertation, including supply chains and system construction.

1.2.1 IT Adoption

Vignette Scenario 1: Consider an organization that has assessed its information system and just acquired a new IT solution that the organization believes will fill a shortfall in its information systems and improve the organization's operations. The new IT system has the functionality to fulfill all organization-defined requirements. Now in the information-systems life-cycle adoption phase, the organization's users start using the new system. The organization sees that the selected IT solution is filling its capability shortfall; however, some in the organization observe that users are not using it to the full extent expected by the organization's management when planning its acquisition. They are not using capabilities that would increase system resilience. As is, the IT solution improves the organization's mission effectiveness, but its unused capabilities could fill the capability shortfall as planned, thus reducing the policy-implementation gap.

Research question 2 ("How can the level of adoption of available IT capabilities contribute to gaps between policy and practice?") relates to organizations' adoption of new IT-driven capabilities. Tactical watch standers afloat use chat the most for command and control (C2) communications, but with less-resilient configurations than they could use. We investigate factors affecting fielded U.S. Navy afloat tactical communications systems' underuse, when using those systems to their fullest would improve mission effectiveness and the resilience of C2 communications. This investigation involves two lines of effort, both exploring the underuse of resilient aspects of afloat tactical networks: chat afloat and communications shift procedures. In each study, we use a mixed methods approach, combining technical system-configuration analysis, survey results collected from system operators, and our analysis of existing policies and procedures. The resulting data help us understand technology-adoption decisions in the field and enable formulating interventions that can close the policy-implementation gap and improve organizational mission effectiveness.

IT Adoption Study 1: Chat Afloat

The U.S. Navy has a policy of deploying "resiliently networked" platforms that can continue operating in communications-degraded or denied environments (CD2Es) [39, p. 11]. However, not all capabilities of these deployed systems are used as intended, enabling the policy-implementation gap between these systems' intended use and that realized. The Navy acquired and deployed chat systems with the resilience-providing capability called a distributed chat architecture to convey C2 communications between ships and shore stations. To discover reasons for the apparent lack of adoption of this resilience-providing capability, we survey afloat command, control, communications, and computers (C4) leadership responsible for configuring and providing chat services to sailors on their ships. Our survey revealed that their reasons include a lack of awareness of the capabilities and benefits of the more resilient configuration. Better technology alone has not made the communications resilient. Understanding how this policy-implementation gap persists can help the Navy improve its use of currently deployed systems and newer systems yet to be deployed. Understanding the users' environment, operations, and motivation are necessary components of successfully deploying and using technical capabilities [40], [41].

IT Adoption Study 2: Communications Shifts

When U.S. Navy ships have a change of operational control (CHOP) between fleet areas of responsibility (AORs), often, they must shift their satellite communications (SATCOM) missions and servicing fleet network operations center (NOC). (We refer to *SATCOM missions* as the satellite resources assigned by the Regional SATCOM Support Centers in the satellite access authorizations (SAAs), which are documents describing the connection parameters. We refer to *SATCOM paths* as the resources used for routing and load-balancing Internet Protocol (IP) traffic. These terms represent two perspectives of the same resources.) If ships have multiple SATCOM paths for connecting their IP network services, they should be resilient to an IP isolation event (i.e., an outage) during the shift. We have observed that the default procedure by which ships shift SATCOM missions and enclave IP services causes them to lose available bandwidth that could otherwise be used during the transition. Having less bandwidth impacts a ship's ability to communicate.

The Navy uses a complex wide-area network (WAN) multiplexing system to interconnect ships and their servicing NOCs ashore. Although sailors necessarily use the new system increments (versions), sailors do not always adopt new procedures to use the designed improvements in function and maximize their limited SATCOM resources during transitions between Fleet AORs. We collected survey data from ships' communications officers during their CHOPs between AORs, which we forecasted from fleet movement schedules. The solicited technical data related to ships' network enclave and SATCOM configurations; the solicited human-factors data related to the way in which the communications officers executed the shift and their satisfaction with the process. With these data, we assessed factors hindering ships from using a procedure optimized for maximizing available bandwidth during shifts, and we developed recommendations for future research to overcome the hurdles to using this system's full capabilities.

1.2.2 Using Standards and Frameworks

Vignette Scenario 2: Wanting to improve its information system security, the organization in our vignette looks to standards and frameworks to better identify and mitigate vulnerabilities. Specifically, the organization seeks to protect the privacy of its people and operations. The standards and frameworks help the organization better identify and mitigate these vulnerabilities; however, their checklist-like appearance lend themselves readily to quick

fixes like assessments (discussed further in Section 2.3) rather than on "achieving a positive outcome for privacy" [42, p. 14]. As such, the organization must take a critical approach to applying the framework, customized to its needs. Further, the framework may contain gaps for some types of threats, which could lead to another policy-implementation gap.

To investigate research question 3 ("How can standards or frameworks guide or hinder organizations' policy implementation?"), we analyzed how well the NIST Privacy Framework [43] can help organizations identify and mitigate privacy inference risks. For organizations that adopt the NIST Privacy Framework, some may expect that the framework will comprehensively approach organization-wide privacy-program management. Although some organizations are pleased with the framework [44], the framework addresses privacy-violating inferences less than organizations may realize. We taxonomized past inference incidents involving a privacy compromise into two categories: re-identification and operational inferences. Then we selected two incidents of each type to analyze against the framework to demonstrate that the framework insufficiently points organizations to inference risks [45].

Our analysis shows how a policy-implementation gap can occur between an organization's intended comprehensive privacy program and the realized privacy program that insufficiently considers inference risks. These findings lead us to recommend improving the framework's coverage of inference risks, especially for the controls in NIST SP 800-53r5 on which the framework relies [46]. Our recommendations relate to the framework's control mappings, integrating concepts related to inference attacks while removing outdated personally identifiable information (PII) dependencies, and the translation of policies into implementation [45].

1.2.3 Enterprise Cybersecurity in the U.S. Government: Policy to Implementation

Vignette Scenario 3: Consider that the organization in our vignette is part of a larger enterprise and is subject to its parent organization's authority to direct cybersecurity policy. Upon receiving a newly directed cybersecurity policy from its parent organization, the organization must determine how to respond. The organization's cybersecurity policy must support its parent organization's directive, external legal and regulatory requirements, and

its own mission requirements. Implementation of enterprise cybersecurity policies therefore vary across enterprises and lack consistency within the policy hierarchy [9], sometimes with organizations' heterogenous noncompliance creating a policy-implementation gap.

To answer research question 4 ("What can cause the implementations of a mandatory cybersecurity policy to vary across U.S. Federal Executive Branch organizations?"), we investigated how various factors affected organizations' implementation of email and Web security policies across an enterprise. The policy asserts it will help protect sensitive information and the reputation of the organization and its components from harm caused by compromising incidents, especially unauthorized impersonation. We found shortfalls in the realized policy implementation across the organizations, yielding a policy-implementation gap.

In this observational study, the policy studied was the DHS BOD 18-01 [19], which applies to about 100 U.S. Federal Executive Branch agencies. Conformance six years after the deadline was surprisingly low with 26 of 97 directed agencies having all their domains conform with BOD 18-01. Through a mixed-methods approach involving direct conformance measurements and a survey of cybersecurity personnel, we found several reasons for non-conformance.

1.2.4 Dissertation Outline

Analyzing the findings of these research studies aided our exploration of the policy-implementation gap. In turn, we propose interventions to help organizations reduce the gap and lay the foundation for future work. In Chapter 2, we explore the policy-implementation gap, the ambiguity-conflict model, and related work. Then we cover our studies in Chapters 3 through 5: "Adopting Resilient Information Technology Afloat," "Using a Commonly Accepted Framework to Protect Against Privacy Inferences," and "Implementing an Enterprise Cybersecurity Policy across the U.S. Federal Executive Branch." We conclude our remarks and summarize our contributions in Chapter 6.

CHAPTER 2:
The Policy-Implementation Gap and Related Work

The policy-implementation gap is the difference between a policy's intended outcome and the outcome achieved by its implementation [10]. A similar concept in cybersecurity is the *semantic gap*, which describes the difference between the concepts communicated in a natural language and how others (including people and computers) interpret them [47]. Studies on the policy-implementation gap reveal conditions in which the gap can occur. Matland synthesizes these studies, building theories on two themes consistently involved in the gap: ambiguity and conflict [9]. In cybersecurity specifically, policy implementations and their validity remain understudied, despite the thriving multi-ten-billion-dollar global market for cybersecurity compliance and risk-management products and services [48]. Cybersecurity studies to date have generally considered how policies can be *verified* (was the policy implemented correctly), classifying failures as configuration or implementation errors [49], or how policies can be *validated* (was the correct policy implemented), classing failures as the policy being insufficient or ineffective [50], [51]. Applying concepts from the ambiguity-conflict model to the policy-implementation gap in cybersecurity policy sheds a new light on cybersecurity policy failure.

A gap implies imperfection. Gunn offers ten preconditions under which an organization can achieve perfect policy implementation [52]. Because some of these preconditions are unrealistic for organizations to achieve, we also use the ambiguity-conflict model [9] to understand how to better reduce the gap.

2.1 "Perfect" Policy Implementation

An early notion of a policy-implementation gap appeared in 1978 in which Gunn offered ten preconditions for "perfect" policy implementation without which gaps occur [52]. We use these to see how various organizational ("agency") factors contribute to a gap. The ten preconditions are:

1. "Circumstances external to the implementing agency do not impose crippling constraints" [52, p. 170].

This describes cases in which the implementing agency does not control these "crippling" circumstances. For example, a storm could cause a multi-day communications outage affecting the availability of an information system.

2. "Adequate time and sufficient resources are made available to [implement the policy]" [52, p. 170].

For government agencies, this precondition is typically satisfied with the first because their resourcing comes from legislation.

3. Not only does precondition #2 hold but also that "the required combination of resources" is available throughout the implementation process [52, p. 170].

In other words, no bottlenecks hinder the flow of the various resources as needed by the policy implementation.

4. "The policy to be implemented is based upon a valid theory of cause and effect" [52, p. 171].

To describe this validity, Gunn quotes Pressman and Wildavsky's view that any policy is a "hypothesis containing initial conditions and predicted consequences[, that is,] if X is done at time t_1, then Y will result at time t_2" [52, p. 171]. Cause X must yield Y as intended or precondition #4 does not hold. This is like a policy's implementation (X) influencing the intended outcome (Y) of protecting against incidents (see Figure 1.3).

5. "The relationship between cause and effect is direct and that there are few, if any, intervening links" [52, p. 171].

Let us define "direct" and "intervening links" with an example. Let P represent the (Web server security) policy that directs a technical configuration $\neg T$ (to disable the deprecated Transport Layer Security (TLS) 1.0 protocol [18]) to achieve goal G (confidentiality and integrity of client connections with the Web server). The organizational policymakers assume that issuing P results directly in $\neg T$'s correct implementation ($P \rightarrow \neg T$) and that by having $\neg T$, the organization directly achieves G ($\neg T \rightarrow G$); therefore, P transitively achieves G ($\therefore P \rightarrow G$). As these preconditions illustrate, however, this organization assumes a lack of "intervening links" (like a possibility to compromise confidentiality by other means) that

would falsify one of these implications. Hudson et al. call this type of agency assumption an "overly optimistic expectation" [10, p. 2].

6. "There is a single implementing agency which need not depend upon other agencies for success or, if other agencies must be involved, that the dependency relationships are minimal in number and importance" [52, p. 172].

As is the nature of preconditions for "perfect" implementation, this could be hard to achieve because organizations have multiple connections to many other organizations and use many software vendors. Implementing a policy across an enterprise's dispersed components can also increase dependency between intra-organizational components [10]. The U.S. Government (USG) tries to overcome the problem of having multiple implementing agencies by placing the responsibility for policy implementation on the head of each federal agency [53], [54].

7. "There is complete understanding of, and agreement upon, the [policy goals] to be achieved; and that these conditions persist throughout the implementation process" [52, p. 173].

In other words, neither misunderstanding nor disagreement among anyone involved with the policy and its implementation can occur. The ambiguity-conflict model recognizes that the policy implementation that minimizes misunderstanding and disagreement occurs under low ambiguity and low conflict, avoiding any suggestion that either of these could be eliminated [9].

8. "In moving towards agreed [policy goals,] it is possible to specify, in complete detail and perfect sequence, the tasks to be performed by each participant" [52, p. 173].

For example, policy ambiguity or conflict in its means of implementation can hinder an organization in specifying implementation tasks [9].

9. The various organizational components involved in policy implementation communicate and coordinate their efforts perfectly well [52, p. 174].

Inadequate coordination can occur, for example, in establishing a common grounding in policy-making [10] and in implementing an information technology (IT) capability across organizational components [41].

10. "Those in authority can demand and obtain perfect obedience" [52, p. 174].

The authority identified here is the organizational leadership directing the policy's implementation. In defining "perfect obedience" to the policy and its implementation tasks (see #8), Gunn quotes Hood: "no resistance to commands at any point in the administrative system" [52, p. 174], which Gunn concedes is not an ideal that we should pursue just to attain "perfect" policy implementation.

Although such a framework is informative, gaining better understanding in how policy implementation can go wrong provides greater value to our analyses. Organizations cannot achieve Gunn's preconditions because of ambiguity and complexity in organizational structure and human factors among management and practitioners.

2.1.1 Organizational Structure

Meyer and Rowan's work on sociological theory describes *decoupling* between organizational policies and goals from practices at the operational level to cope with ambiguity and complexity in organizational structure. They argue how "organizations that reflect institutional rules tend to buffer their formal structures from the uncertainties of technical activities by becoming loosely coupled" [55, p. 341]. This decoupling varies in form and cause across diverse application domains, including artificial intelligence (AI) ethics [56], post-communist bureaucracies [57], and social movements [58]. These examples show the consistent paths by which the substantive goals espoused by organizations' stated policies separate from practices by and in their members' day-to-day work. As Gunn identifies in preconditions #6, #7, and #9, such decoupling contributes to the policy-implementation gap [52].

Relating organizational structure with technologies is common in the fields of safety science, medicine, and systems engineering as the relationship affects policy outcomes in complex engineered systems [59]. We believe similar gains in cybersecurity are possible by understanding the social actions surrounding cybersecurity decisions [60]. It is a special case of

the broader study across fields of the relationship between organizational goals operationalized as policies and the behaviors of organizations as driven by the day-to-day work of the people comprising them. Ethnographic work suggests that information-security workers are sometimes unconvinced that policy implementation improves cybersecurity [61].

2.1.2 Human Factors Among Management and Practitioners

An organization's management must understand reasonably well the organization's cybersecurity policy implementation from a sociotechnical perspective and associated risks to structure the organization for effective cybersecurity [62], [63]. Such understanding supports management's obligation to certify an organization's cybersecurity program. Much like safety-risk analysis, organizational management must assess, and sometimes assert, the effectiveness of its cybersecurity controls in achieving the desired goals [64]. To assess management's capacity for cybersecurity assessment, the business management consultant firm Ernst & Young surveyed more than 1,000 organizational cybersecurity leaders [65]. They found that from 2020 to 2021, the percentage of surveyed organizations' boards of directors who were "extremely confident" in their organizations' cybersecurity risk mitigation measures decreased from 20% to 9%, and that the number of boards of directors discussing cybersecurity at least quarterly increased from 29% to 39%. In a separate study, organizations having a board-level technology committee had better cybersecurity oversight and, as the technology committee matured, lowered their likelihood of breach compared with organizations without the same committee [66]. Bringing management and cybersecurity practitioners closer together in understanding should help resolve ambiguities and conflict while reducing the policy-implementation gap.

Understanding human factors in cybersecurity-policy implementation can help learn more about the policy-implementation gap, as they relate to Gunn's preconditions #4, #5, #7, #9, and #10 [52]. A thriving community of human factors research operates under the banner of "usable privacy and security," but its focus is "the design, construction, and deployment of systems that people can use to secure computers and personal information" [67, p. 7], [68]. For example, incentives and individual work practices play heavily into human factors [31], but vary across types of organizations, like government and non-government [69]. Researching the complexities of human interaction with organizational information systems can broaden the focus of usable security from individual psychology to phenomenology

at the social and organizational scale. These efforts help us bridge the divide between the technological problems of meeting security-policy implementation requirements and the managerial problems of establishing policies that serve security goals when implemented.

2.2 The Ambiguity-Conflict Model of Policy Implementation

Work following Gunn approaches the policy-implementation gap from a more realistic perspective that embraces the complexity inherent in the subject and works through the implementation gaps carefully [10]. For example, Matland offers the ambiguity-complexity model of policy implementation [9], summarized in Figure 2.1. It describes the key factors that influence organizational-policy implementation. Ambiguity and conflict can each arise about a policy's goals (at the *macroimplementation* level) and its implementation tasks (at the *microimplementation* level). Ambiguity is not necessarily bad because it can reduce conflict (because fewer details exist over which to argue) and permit greater autonomy in implementation according to local needs.

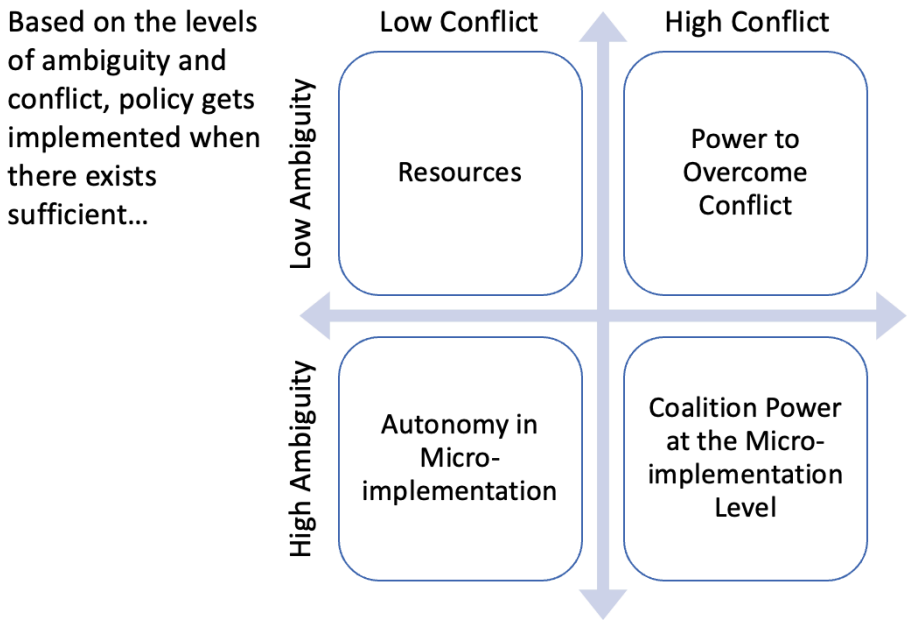


Figure 2.1. The Ambiguity-Conflict Matrix. Adapted from [9].

Based on the levels of ambiguity and conflict, the model suggests which factors have "the greatest influence on the implementation outcome" [9, p. 160]. When both ambiguity and conflict are low (called *administrative implementation*), we expect that implementation will occur once sufficient resources are available. In a sense, cybersecurity policies' technical implementations should be administrative because of the uncontentious goal of protecting against cybersecurity incidents and the wide availability of tools to help system administrators with cybersecurity management. As ambiguity about a policy's goals or implementation tasks increases, practitioners need greater autonomy to use their professional judgment in implementing the policy. Instead of increasing ambiguity, if conflict increases, the entity having the greater organizational power will decide how to resolve the conflict about the policy's goals or implementation tasks. When both ambiguity and conflict are high, those who control the resources and those with the ability to implement the policy must agree on how to proceed; that is, they must form a coalition to implement the policy.

Matland's synthesis of policy implementation literature suggests two primary implementation perspectives, top-down and bottom-up [9]. Matland observed that top-down policy implementation approaches are best suited for implementing unambiguous policies and bottom-up approaches are better suited for ambiguous policies in which local practitioners must customize their implementation to achieve the intended policy outcome [9, pp. 170–171]. Proponents of top-down policy implementation tend to focus on the administrative aspects of implementation while minimizing political influences and human factors. In contrast, proponents of bottom-up policy implementation tend to focus on the factors that affect how practitioners implement policy, including their autonomy, relevant skills, policy interpretation, and position within the organizational structure. In both cases, local practitioners must have some latitude to apply policy according to local needs.

The ambiguity-conflict model helps for assessing policy implementation in diverse public-policy situations but has not yet been used for analyzing organizational cybersecurity-policy implementation. For example, the model helped discover reasons for the policy-implementation gap in the implementation of the Vanguard healthcare model in the United Kingdom [70]. In this case, centralized policymakers issued a high-ambiguity, low-conflict policy, intending to learn from the local-office implementations and build a framework for guiding implementations at other offices. The bottom-up approach did not yield the intended national framework because widespread microimplementation enabled implementations

too diverse for policymakers to find generalizable patterns for forming a framework. As a different example, the model helped discover reasons for the policy-implementation gap in the U.S. Federal policy responses to a natural disaster, hurricane Harvey in 2017 [71]. Zarb and Taylor select four such policies, one that fits into each quadrant of Figure 2.1. In this case, ambiguity and conflict about a policy's goals proved to be greater predictors of the policy-implementation gap than ambiguity or conflict about a policy's means of implementation.

Matland describes the implementation of policies with high ambiguity and high conflict as having a *symbolic* implementation [9]. Symbolic implementation, sometimes called gratuitous implementation [72], is like going through the motions of implementation to appease a compliance requirement, appearing to comply but not necessarily complying with the policy [73]. It encourages a large policy-implementation gap because of the policy's limited implementation. Symbolic implementation relates to the idea of *managerialization*, in which the goals of corporate management, rather than the substance claimed by the language in the policies, colors its interpretation of legal and other policy requirements [33].

The ambiguity-conflict model shares some characteristics with neo-institutionalism [74]. Skille and Stenling appears to provide the only research before ours that connects the ambiguity-conflict model with neo-institutionalism [75]. Specifically, in autonomous microimplementation, various imperatives force people to adapt policy; however, all implementations then derogate from policy, thus separating practice from the policymakers' intent. These microimplementation practices further decouple an organization's goals for a policy and the policy's implementation [55], widening the policy-implementation gap. For example, practitioners must translate and adapt cybersecurity standards and frameworks to the local and unique needs of the information system [76]. Even having standardized computing systems does not avoid decoupling because information systems are sociotechnical.

2.3 Reducing Policy-Implementation Gaps

A variety of factors can cause policy-implementation gaps, including focusing too much on components, human error, or compliance. Understanding people and how they cooperate in organizations with information systems (applying social and structural considerations) can help us better understand and reduce the policy-implementation gap.

A policy that tries to make a system reliable by making the system's components reliable is unlikely to succeed because reliable components do not imply reliable systems [59], [77]. Failure at the system level can occur without component failure, so the implementation fails to achieve the policy goal of having a reliable system. For example, in the nuclear-power accident at Three Mile Island, great system complexity, tight coupling, inconvenient control-room design, and insufficient operator procedures, training, and education all contributed to the accident [77]. None of these can serve as an independent cause of the gap in the implementation failing to achieve the policy's no-nuclear-accident goal. Although formal investigation boards cite human error as the root cause, operators were not positioned to avoid the accident because their ability to take actions or make decisions to avoid the accident had structural limitations [77]. Human error is often the cause of last resort for system failures because humans are imagined to be "in control" of the systems; however, the very complexity and structure of the systems can prevent humans from truly controlling them [78], [79], [80].

Cyber incidents are often likewise complex, involving flaws in complex system configurations and sometimes operator error [81], [82]. Like the accident on Three Mile Island, operators (users) can be unfairly blamed for cybersecurity incidents, which is unhelpful for finding the causes of the policy-implementation gap [63]. Many other factors can contribute to cybersecurity incidents, including counterintuitive or insecure software design [83], [84], [85], unreasonable expectations of users [79], [86], ineffective or incorrect advice from cybersecurity experts [87], and environmental and social pressures on users [60], [88].

Compliance to security policy requirements as a checklist can supplant the goal of system-level security, as if compliance were the goal rather than security. Sometimes a checklist is helpful; for example, a simple 7-item surgical operations checklist covering straightforward items, including the planned procedure and the names and roles of participants in the surgery, improved the quality of surgery on a global scale [89]. A framework or checklist can guide compliance, risk perception, and management tasks [90]; but following a checklist is insufficient for achieving cybersecurity [50], [91], [92]. One may find this counterintuitive because we can represent checklists logically as conjunctive predicates (like Equation 1.1) and implement them effectively as finite-state machines. However, as creating a checklist on which surgeons can rely completely for surgeries is intractable, so would cybersecurity professionals in creating a comprehensive cybersecurity checklist [93]. As such, cyberse-

curity professionals often adopt a *risk-based* approach, as in safety assessment [38], [59], [94].

Treating checklists as dispositive claims on appropriate security measures mistakes the checklists for code. Checklists communicate risks or minimum performance standards, not computer-executable instructions [93]; therefore, relying solely on checklists can lead to symbolic implementation. Instead, applying professional judgment to interpret the goal of the checklist and each item as intended by the checklist's authors is more effective [91]. If we could specify exact requirements for achieving specific levels of cybersecurity as computer-executable instructions, verification of the implementation would avoid this error [92].

2.3.1 Social Considerations

Applying sociology to cybersecurity adapts the research methods of social and organizational sciences and safety engineering to better understand the role of organizational decision-making in the performance of engineered artifacts and computing systems [14], [95]. Many are interested in how human factors affect security outcomes [67], [96], [97]. For example, Haney and Lutters champion the concept of cybersecurity advocates in organizations [98], like Beris et al.'s "security champions" [82]. Safety engineering has seen large gains from applying the methods of social science and organizational and human factors analyses [59], [95]. Cybersecurity can also benefit by applying these methods to learn more about the policy-implementation gap [99].

Concepts from these methods, like a *compliance budget* and organizational security culture, help us better understand how information-system users affect the policy-implementation gap. Beautement et al. offer the idea of a compliance budget for organizations to manage when considering and implementing cybersecurity controls [100]. It balances the measure of employees' efforts spent on compliance with their perceived cybersecurity benefit. Once the budget is exhausted, employees become more likely to find paths of less resistance to accomplish their work, thus circumventing the burdensome cybersecurity controls and widening the policy-implementation gap. If one quantifies the compliance budget, predicting how the size of the policy-implementation gap changes when the budget is expended may become possible. As a hedge against such circumvention of controls, organizations can shape perceptions of the measure of compliance efforts and their cybersecurity ben-

efit [100]. Beris et al. build on this concept by recognizing the heterogeneity of people's views toward cybersecurity and offering a taxonomy to influence organizational cybersecurity culture [82]. Grading employees' emotional position toward cybersecurity and their understanding of cybersecurity risk can inform organizational interventions to shape employees' cybersecurity behaviors [81]. Objective measurement remains a challenge in these methods.

2.3.2 Structural Considerations

Technology creates a structure for the world that shapes what is possible and affords what choices people can make. Some of these choices might be to change the technology so technology is simultaneously the context for and the product of choices [101]. Reidenberg considers government regulations on network and communications technologies, arguing that government regulations must consider cyberspace's "Lex Informatica" for a more effective regulatory approach to the cyber domain [102]. We extend this concept by claiming that organizational policymakers must consider their cybersecurity practitioners' systemic-structural context in policy formation. With his constraints model, Lessig extends Reidenberg's concept, agreeing that code becomes the policy of the cyber domain (i.e., the "Lex Informatica") because code, rather than government regulation, defines the architecture and its own practical function [103]. In cyberspace, code forms the environment's enabling and limiting characteristics, that is, the structure. Reidenberg and Lessig both argue, however, that choices in the design of human artifacts simultaneously control social structures while being a product of social structures; therefore, designing these artifacts requires expertise to understand their interactions with social norms. Norms are the real-world social structures that define culturally accepted interactions [103]. In other words, the most effective cybersecurity policymakers and practitioners form a symbiotic relationship of understanding how their design and implementation choices affect users' ability to operate, which is necessary for managing sociotechnical systems [104], [105].

Competing views on controlling complex sociotechnical environments to decrease the risk of failure come from Perrow [77], Levison [59], and Vaughan [106], [107]. Perrow argues that we cannot engineer safety into the designs and operations of complex systems with tight coupling between causes and effects, like nuclear power plants [77]. In these cases, he recommends society avoid their development and proliferation. As the coupling between

causes and effects loosens, even when keeping high complexity, greater tolerances (in time and in sequence of events) emerge for intervening between the causes and effects. On the other hand, Leveson claims that organizations can engineer safety into very complex systems, in part, through Systems Theoretic Process Analysis (STPA) [59]. One can use STPA for designing safety into systems during the system design process, like the methods used for designing cybersecurity into systems [53]. Vaughan emphasizes that the relationship between the systemic-structural context of the individual, routinization of non-conformance, and organizational compliance decisions can lead to the emergence (in a systems-theoretic sense) of disastrous outcomes (like the Challenger space shuttle launch decision) [106], [107]. Cyber systems are complex with tight coupling between causes and effects in some aspects (as with race conditions and the speed at which computers operate compared to users) but loose coupling in others. Examples of loose coupling in cybersecurity include how interfaces are not always clear [67], [108] and people differ in training and education [82]. Despite their differences, Perrow, Leveson, and Vaughan examine catastrophic accidents through similar analyses of system structure to highlight the factors affecting the accidents.

CHAPTER 3:
Adopting Resilient Information Technology Afloat

Portions of this chapter appear as "Acquired and Deployed but Not Adopted: Lost Mission Effectiveness without Resilient Chat Afloat" by C. B. Landis and J. A. Kroll in the *Proceedings of the Twenty-first Annual Acquisition Research Symposium*, 2024, vol. 3, pp. 199–214. Available: https://dair.nps.edu/handle/123456789/5161 [41].

U.S. Navy sailors continuously operate communications systems that are critical to modern command and control (C2). The Navy has designed and deployed these communications systems to maximize its resilience in communications-degraded or denied environments (CD2Es). As a top-down policy-implementation approach, it has been only partially effective, however. This study reveals that cases exist of sailors underusing some of these systems' resilience-providing capabilities. Such underuse yields a policy-implementation gap between the capabilities intended by Navy leadership and the capabilities realized in the fleet. In exploring the research question, "How can the level of adoption of available information technology (IT) capabilities contribute to gaps between policy and practice?," we study two underuse cases of U.S. Navy communications systems.

The Navy relies on its afloat Internet Protocol (IP) tactical networks for conveying mission-critical data between ships and shore facilities. Although these system capabilities provide specific functionality, the desired outcome, resilience, is a non-functional, system-level property, like security or safety. Such non-functional properties are easily proven only in one direction, that is, when they fail [27], [109]. Resilience is invisible yet critical.

We offer an interdisciplinary approach in two studies to investigating the organizational and technical hurdles to using resilience-providing capabilities in deployed communications systems. Because the author's fleet experience provides the initial motivation for this study, part of our investigation is to discover whether these observations are isolated or more widespread. To explore the situation in each study, we survey a sample of sailors responsible for managing and using the subject systems and analyze system configurations to devise interventions for improving these systems' resilient configuration and use. If we fail to understand why these sailors underuse deployed systems, we risk similar underuse in future

capabilities and, ultimately, lost mission effectiveness. Our two studies are on chat afloat and communications shifts.

The Navy routinely conveys real-time C2 communications between tactical watch stations across multiple ships and ashore by chat, which operates over IP links. As such, the Navy has long required that chat and IP capabilities be resilient [110]. To be resilient, these systems should gracefully degrade when links in an IP path disconnect, continuing to provide communications paths between those nodes that remain connected. A recommendation coming out of the U.S. Navy exercise Trident Warrior '04 to use *distributed chat architectures* (explained in Section 3.2.1) subsequently became a Navy requirement and was tested satisfactorily during Trident Warrior '05 [111]. This represents bottom-up feedback for distributed chat architecture capability becoming a top-down policy implementation. Other studies on chat use afloat focus on technical aspects [112], user interaction with the chat software [113], or both [114], [115] and all date from between 17 and 20 years before our work, which provides a current perspective. In this study, nearly all our survey respondents report that they do not use afloat chat servers for providing chat services to sailors aboard their ships. Instead of configuring distributed chat architectures, they routinely use configurations that are intolerant of wide-area network (WAN) interruptions, revealing a decoupling of the implementation from the Navy's goals, a policy-implementation gap.

While deployed, ships transit from their home ports to the maritime destinations their duties and assigned missions take them, conducting communications shifts along the way. Communications shifts typically occur when these transits involve crossing boundaries from one fleet commander's area of responsibility (AOR) to another fleet commander's AOR, an event called a change of operational control (CHOP). The Navy's WAN architecture connects ships' multiple networks over multiple satellite communications (SATCOM) paths with other ships and fleet network operations centers (NOCs) ashore [116], as shown in Figure 3.1. Recent studies on improving this WAN architecture focus on technical aspects but not on network administrators' use of them [117], [118], [119], making our study unique in its sociotechnical approach. In this study, ships' cross-AOR communications shifts often entail unplanned outages and troubleshooting, even when a ship's multiple SATCOM paths should provide the redundancy to be resilient to complete IP connectivity outages.

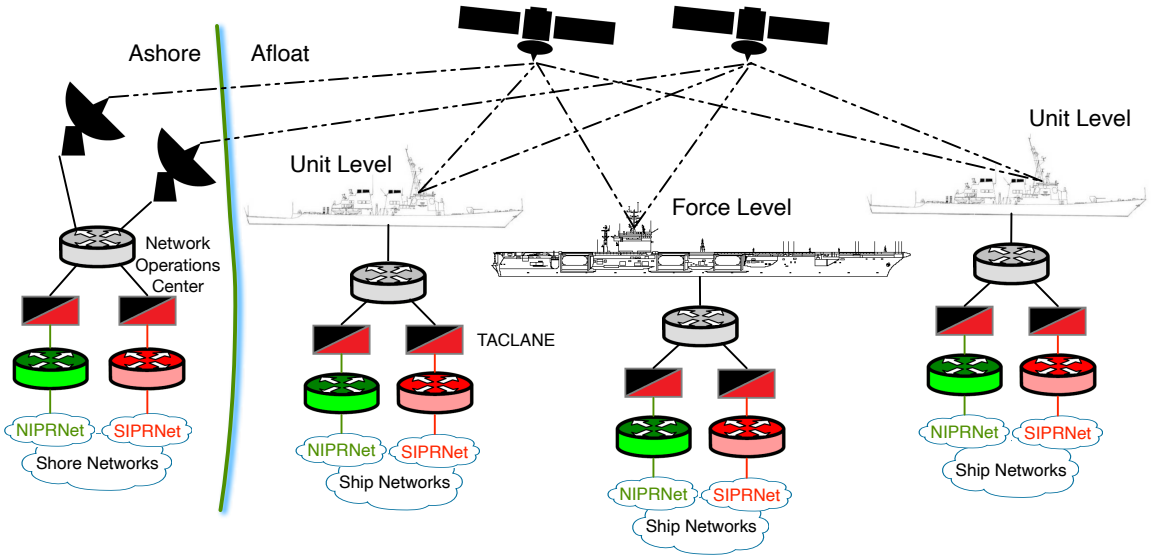


Figure 3.1. Conceptual Overview of Fleet Connectivity. The Navy primarily sends unclassified communications over the Non-secure Internet Protocol Router Network (NIPRNet) and command and control (C2) communications over the Secure Internet Protocol Router Network (SIPRNet) [114]. Adapted from [116].

To discover what is happening at the microimplementation level, we take a two-prong investigative approach to how a sample of sailors manage and use these systems for conveying C2 communications. We administer a survey of current and former afloat command, control, communications, and computers (C4) leadership (those responsible for C2 communications systems) as part of our analysis of chat use. We also administer a separate survey of ship's communications officers during their CHOPs to record quantitative and qualitative data about their SATCOM and IP network communications shift experience. We use the ambiguity-conflict model (introduced in Section 2.2) to recognize factors that widen the policy-implementation gaps in both cases [9]. Overcoming the hurdles that these sailors faced to using resilient-providing capabilities in deployed communications systems furthers the concept of assured C2 in distributed maritime operations, thus improving warfighting effectiveness.

We review the theoretical context in Section 3.1 before addressing and analyzing our survey data for afloat chat use and communications shifts in Section 3.2 and Section 3.3,

respectively. We discuss our findings, explain devised interventions, and conclude our remarks in Section 3.4.

3.1 Background

Understanding why people are not using a technology to its fullest capabilities requires a sociotechnical examination, approaching "society and technology together as one coherent system" [104, p. 1]. To apply Matland's ambiguity-conflict model of policy implementation to organizational policies on IT adoption, we consider how the structure of organizations and their technologies affect each other. In this section, we introduce background and related work about the Navy's deployment of computing capabilities afloat, the effects of automation, the status-quo bias, and how the concept of technology-use mediation integrates with the ambiguity-conflict model.

3.1.1 Naval Technology Acquisitions

The Navy establishes programs of record (PORs) for managing C2-supporting communications systems throughout their information system life cycles. A POR is an organization with the task of acquiring or building, integrating, deploying, maintaining, and discontinuing a system. The Office of the Chief of Naval Operations (OPNAV) defines requirements and funds PORs' work (through systems commands (SYSCOMs) like Naval Information Warfare Systems Command (NAVWAR)) whose products get deployed on Navy ships for sailors to use. Because PORs necessarily focus on meeting their sponsored requirements, inter-POR coordination easily becomes a secondary effort. However, because resilience is a system-level property, it cannot be described concisely as a requirement and its implementation requires action across organizational functions. Communications systems' interconnectedness needs cooperation between PORs during their development to ensure their resilience when deployed. Once deployed, improving their resilience is challenging because it requires coordinating sailors' innovation efforts across multiple ships. As a person in a position to oversee multiple PORs, a NAVWAR commander desires greater inter-POR cooperation to achieve greater system integration [120]. Type commanders (TYCOMs), like Commander, Naval Surface Forces (CNSF), collect requirements for new systems or system improvements from sailors in the fleet using current systems. In turn, fleet commanders consolidate TYCOM requirements and submit them to the OPNAV for consideration and

funding, as shown in Figure 3.2. But, as engineering systems professionals say, technology is only part of the solution [121]. Even when PORs cooperatively provision all the technical aspects well, people still must operate the system proficiently for it to yield its benefit to the Navy's mission.

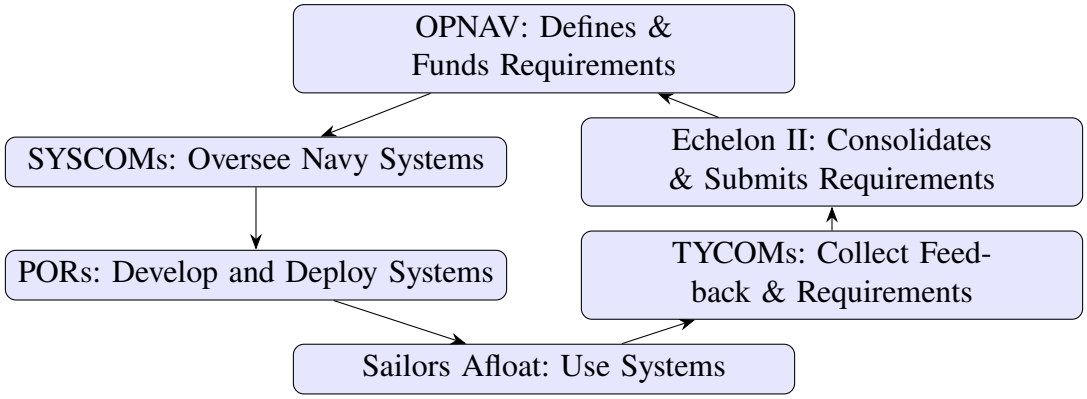


Figure 3.2. A Summarized Diagram of How the Navy Defines Requirements and Deploys Systems

When introducing the Naval Tactical Data System (NTDS) (a combat information system for fusing sensor data, displaying a common tactical picture (CTP) across multiple ships, and enabling weapons pairing and firing), the Navy trained sailors to be the system experts rather than relying on vendor technical support [122, p. 254]. This enabled sailors to use the system effectively across multiple ships, especially because of their combined knowledge of the system, its operating environment, and the culture influencing its users. Vendors and POR personnel rarely deploy to operate the deployed systems that they developed but may visit a deployed ship to fix a system that is broken beyond sailors' repair capability. Today, being system-of-systems experts of communications systems is important for sailors to maximize their use of deployed information systems.

3.1.2 Theories of Operator Responsibility

The theory of the ironies of automation suggests that a barrier to this systems-level knowledge exists in the form of system automation, that is, a system works when operating

according to POR baseline configurations and guidance so little *perceived* operational need to dig deeper into how the system actually works exists [86], [109], [123]. Although system details are available in the system itself, these details are not in any POR-sponsored training because it is unnecessary for the system's baseline operation, which only meets the OPNAV requirements [124]. However, one can discover these details by teaching oneself, tracing cables and studying network diagrams, router configurations, and ships' traffic flows [116].

Bainbridge [86] and Sullivan [124] find that greater automation in systems does not reduce the role of the human operator but simultaneously alters it and makes it more critical to system behavior. For example: "Because the routine operation of ADNS [(Automated Digital Network System)] requires little operator intervention[,] the tendency might be to take a more hands off approach to communications management when in fact the opposite is true" [124, p. 35]. In other words, the operators should take a more hands-on approach to understand the role of the automation. That greater automation tends to falsely reduce users' responsibilities persists as automation and system complexity increase [79], [123]. In other words, the lack of knowledge but burden to acquire it combined with the lack of widespread perceived need for knowledge on the inner workings of afloat communications systems, like ADNS, could block casual attempts to improve their use.

Another theory that can affect technology capability nonuse is the status quo bias, especially when uncertainty in the outcome exists [125]. When people lack the information explaining why a policy change or new technique is being introduced, they are more likely to stick to their current practices than adopt the new technology [126]. If the explanatory information is available but difficult to find, the inconvenience of searching for and processing it increases their status quo bias. Increased automation in a system can hide system functionality from its users, thus discouraging users from learning more about how to use it better, which engenders users' status quo bias.

A question with any new system is whether people should adapt the system to existing processes or adapt existing processes to the system. For example, in adopting a new enterprise resource planning (ERP) tool, Apple first tried to adapt the ERP tool to its existing workflows. When Apple Chief Executive Officer (CEO) Steve Jobs heard that this effort had become too expensive and cumbersome, he took a risk, directing Apple to cut its losses and start over, this time adapting its workflows to the new ERP tool, anticipating correctly that Apple

would find greater success [127]. Orlikowski et al. refer to the technology framing activity that Steve Jobs did as *metastructuring*. That is, to "deliberately adapt computer-mediated communication technologies and their use to particular contexts and change those contexts to accommodate use of the technology" [128, p. 424]. The Navy has adapted network routing protocols and chat software ("computer-mediated communication technologies") to the Navy's specific needs for conveying C2 communications (a "particular context"). Further, the Navy has adapted the contexts on Navy ships to use these technologies to improve ships' ability to accomplish their missions. These adaptation actions constitute metastructuring. An organization's metastructuring approach to adopting new IT affects the organization's success.

Matland's ambiguity-conflict model supports top-down and bottom-up policy implementations [9]. In IT adoption, the organization's policy is to adopt the technology to improve its mission effectiveness and help the organization to achieve its goals. The Apple ERP metastructuring example provides a top-down approach to an IT adoption policy implementation. Practitioners also take part in a technology's metastructuring, which can affect the technology's microimplementation and influence the organization from the bottom-up [129]. For example, a study on radiology departments adopting computed tomography (CT) scanners (a top-down implementation) found practitioners' responses to and interactions with the new technology profoundly affected department reorganizations (bottom-up metastructuring) to better include the technology into work processes [130].

In influencing how an organization use a technology, Orlikowski et al. study a type of metastructuring called *technology-use mediation*. They define it as:

> Deliberate, ongoing, and organizationally-sanctioned intervention within the context of use that helps to adapt a new communication technology to that context, modifies the context as appropriate to accommodate use of the technology, and facilitates the ongoing effectiveness of the technology over time. [128, p. 424]

In their study, they identified the network administrators as the mediators because of how they intervened to influence the technology's use. Playing the roles of "champions, trainers, and local experts" [128, p. 437], mediators simultaneously structured technology to enable use and interacted with other users to understand use requirements to improve their structuring

activities and system use. These mediators helped influence the technology's use, ostensibly because of their rapport with the users and their organization's supportive position toward their mediation efforts.

3.2 Perceptions of Chat Servers Afloat

In the technology-use mediation context, the Information Professional Officers are in position to be mediators but do not act as mediators. Information Professional Officers fill most C4 leadership billets, so we solicited survey responses from them using the survey instrument shown in Appendix A. The Information Professional Officers are responsible for the afloat server configurations to enable a distributed chat architecture. Also, just as the network administrators in Orlikowski et al.'s study are newsgroup users [128], many Information Professional Officers use chat at their watch stations. This proximity to users enables them to build rapport with other users, understand users' needs and constraints first-hand, and influence users' technology use.

With installations of the current afloat network server suite architecture starting in 2013, the population of Information Professional Officers for this survey was about 510. This assumes a ten-year linear growth in installations to 170 ships, synchronized biennial officer transfers, and an average of 1 Information Professional Officer per ship, some ships having none and others several. With 6.25% annual attrition [131], we estimated that approximately 435 of the 510 officers were available to survey in 2023. We collected 69 survey responses (provided in Supplemental 1) from those who have served on a ship at the force level (33) or unit level (30) (6 skipped this question).

We have taken steps to reduce survey sampling and participants' self-selection bias. We used multiple recruitment channels, including the Navy Information Professionals milSuite collaboration page that the Information Professional Community leaders use for issuing official community information [132] and the Information Professional Team on the Navy's FlankSpeed [133]. Our population selection is "purposive" in that we solicit participation from Information Professional Officers rather than any sailor that uses chat afloat [134]. Most chat users are unaffiliated with network administration. Our participants' ship class demographic accounts for a nearly even balance of those on force-level ships that would host group-level chat channels and unit-level ships that are less likely host such channels.

In summary, our results may not represent the population but still provide valuable insights into our research question.

In Section 3.2.1, we explain and contrast the more-resilient distributed chat architecture and the afloat chat server configuration that our survey data reveal is the norm. We describe our survey results for the respondents that do not use the afloat chat servers and those who do in Sections 3.2.2 and 3.2.3, respectively. The responses help us discover the organizational and technical factors hindering these sailors from using their ships' chat servers, which we discuss in Section 3.2.4.

3.2.1 Chat Afloat

When configured as a *distributed chat architecture*, shown in Figure 3.3, all afloat users connect their chat clients to their ship's chat server, which the ship network administrators have federated with the servers on other ships and in the fleet commander's maritime operations center (MOC) ashore. This is the ideal configuration to achieve a resilient configuration. Federating servers enables the *bridging* of chat channels across servers. A chat channel owner (i.e., fleet commander, strike group commander, etc.) can host the channel locally and bridge it across federated servers. If a ship loses connectivity with the shore, all shipboard users stay connected in the chat channels locally and can continue communicating with each other. While disconnected from the NOC ashore (through which ships connect to resources ashore, including the MOC), if the ship remains connected with other ships through an IP path, these ships' chat servers can remain federated. In other words, the chat users on multiple ships can stay connected through chat, even without connectivity with the MOC's chat server. Upon reconnecting with the shore, the channels automatically resynchronize the last pre-configured number of hours of history. Thus, this distributed chat architecture provides resilience for a transient shoreless environment.

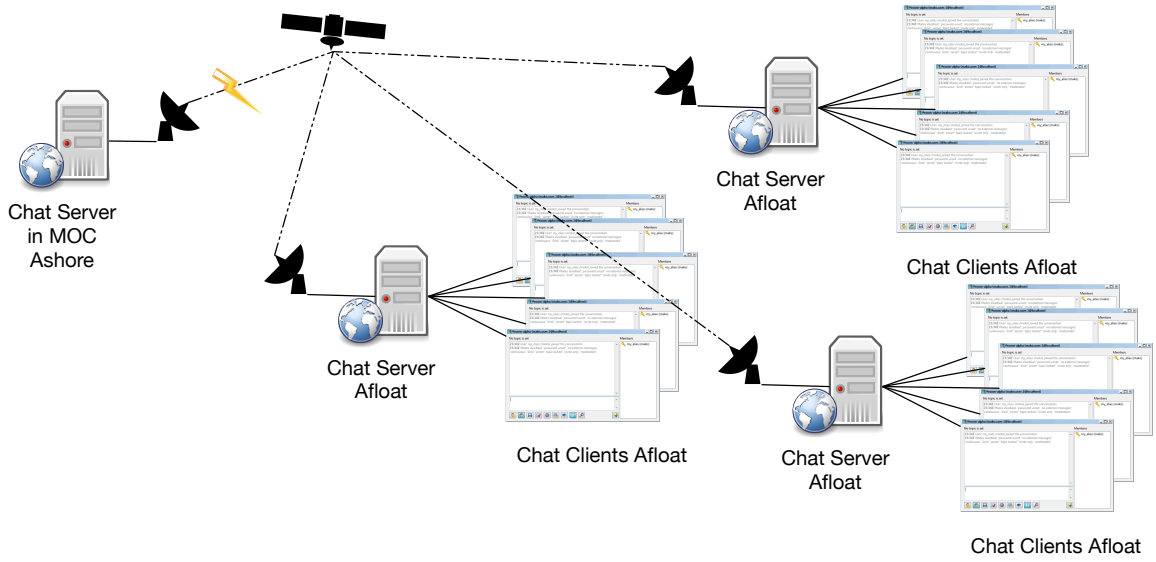


Figure 3.3. Conceptual Overview of a Distributed Chat Architecture with Clients Connecting to Local Servers and Servers Federated Across Units

When logging in to their chat clients, users specify by IP address the server(s) to which they want to connect. Despite having chat servers on ships, users typically connect their chat clients only to servers ashore, as shown in Figure 3.4. This configuration is intolerant of network interruptions (i.e., all clients drop upon losing connectivity with the shore) and incurs additional bandwidth demand on SATCOM paths with many afloat clients connecting ashore instead of a single server. Even if system administrators federate afloat chat servers with off-ship servers, users still might not configure their chat clients to use their ship's chat server. Because enabling shoreless chat by a distributed architecture depends largely on afloat chat server administrators and end users' client configurations, we must understand the network administrators and users better.

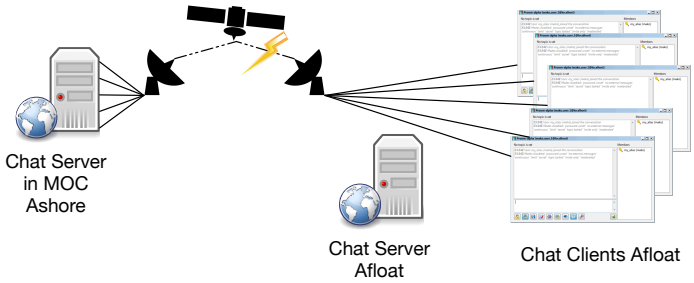


Figure 3.4. Conceptual Overview of How the Navy Uses Chat Afloat, Bypassing the Afloat Chat Servers

All ships had an Internet Relay Chat (IRC) server that could participate in a distributed chat architecture until April 2021 [135], [136] when an update of a third-party dependency broke the chat federation capability [137]. Most survey respondents (45) report starting their afloat tour before the POR stopped supporting the chat server federation capability, so they could have used it. Losing the chat federation capability did not cause enough protest to demand an immediate replacement to provide that capability, as is clear from the lack of a replacement and indicative of its nonuse. No survey participant identifies the loss of federation capability as a reason for why they do not use the capability. The POR for ships' computer network suite is working on testing and fielding solutions, including in the next version of ships' network services suite and with other chat software to integrate with newer communications capabilities. Any how-to sever configuration guide for afloat network administrators or client user guide describing a distributed chat architecture must therefore cover newer chat software solutions that are still in development.

3.2.2 Hurdles Inhibiting Afloat Chat Server Use

As part of providing the chat capability, our survey respondents direct and oversee the afloat chat server configuration. If they had not configured it to operate in a distributed chat architecture, then no user on their ship could connect to the afloat chat server to participate in a distributed chat architecture. The reasons for our survey respondents not providing the afloat chat server capabilities to sailors aboard their ships vary, as shown in Table 3.1. (The row totals vary due to the branching shown in Figure A.1.) Twenty of 69 participants are unaware of afloat chat servers. Of the 44 of 69 that are aware of afloat chat servers' existence, 32 do not use them at all. They are unaware of the benefits of their use and untrained on

their capabilities. For example, four Combat Systems Information Officers (CSIOs) and a command, control, communications, computers, and intelligence (C4I) Officer (participants P7, P34, P42, P45, and P55) report an apparent lack of benefit, including not knowing about the bridging capability. "I wouldn't want to add another chat capability separate of the existing C2 chat rooms used by FLTCDRs [(fleet commanders)] and units. I'm not sure what purposes this chat server will serve" (P7). "Even when I educated staff members about it, they wouldn't use it. The customers didn't know of bridging chat rooms to site [*sic*] that as a requirement nor did any of the ITs [(Information Systems Technicians)] onboard know how to do that" (P34). For every nonuse reason, except "I don't know," between 67% and 100% of participants could have used the chat federation capability. Transitively, sailors aboard these survey respondents' ships were not using the afloat chat server because the service was not provided, and it was not provided for the reasons listed here.

Table 3.1. Reasons for Not Providing the Afloat Chat Server Capability, Corresponding with the 69 Responses to the Survey Instrument in Appendix A

Factor	**Yes**	**No**	**Don't Know**	**Skip**	**Total**
Aware of Server's Existence?	44	20	-	5	69
(Aware) ↪ Ship's Sailors Use Server?	9	32	2	1	44
(Not Use) ↪ Why not Use?	-	-	5	1	32
Unaware of Channel Bridging Capability					15/32
Unaware of the Benefits of Its Use					11/32
No Time to Receive or Provide Training on Its Use					9/32
Handling Too Many Other Network Problems					7/32
Unaware of How to Connect					4/32
No Operational Need (in Shipyard)					1/32

Ten survey participants (P3, P16, P19, P32, P35, P38, P46, P62, P63, and P67) claim that if they had known about the server, they would have used it for improving communications

internal to the ship. For example, P32 claims the use would be "To allow for the internal watch stations on the ship to maintain text comms even when down IP services." As an alternative use, P67 claims, "I would have liked to be able to chat to personnel aboard the ship internally. Sometimes emails are too slow especially if I know they on the computer."

Responses to open-ended questions in our survey instrument reveal qualitative narratives for the sources of nonuse. Some participants' survey responses indicate their nonuse of afloat chat servers relate to training requirements and culture, as described below.

Training Requirements

Requirements supporting the Navy's chat system likely lack a champion because the Navy treats the chat application as a capability of the afloat computer network suite POR rather than its own POR. Participant P29 claims that afloat network administrators do not fully understand the *operational* capabilities of ships' networks, including the chat server, because the Navy provides no formal introduction to them. To make the point, P29 uses the example of the SharePoint server, which is part of the network installation but is neither configured nor introduced to the administrators responsible for maintaining the network and its servers. But for users to get the most out of the tool in its operational context, administrators must configure the system relative to that context, whether that be federating chat servers or creating appropriate user groups as SharePoint site owners. P45 concurs: "Likely wouldn't have used it, as it would have likely not been configured and sailors [(network administrators)] are not adequately trained to configure these servers without guidance." A similar Government Accountability Office finding involving IT nonuse in the U.S. Immigration and Naturalization Service blamed a lack of training [138]. This is like the lack of information transparency described in Section 3.1 in that administrators do not have the requisite information to configure and operate chat afloat resiliently. In other words, the surveyed Information Professional Officers were not empowered to be technology-use mediators, despite their positioning in the organization to intervene.

Training courses are a type of operational need; a course for chat does not exist. The OPNAV determines operational needs and defines Navy requirements to meet these needs. All the Navy training courses for afloat computer network administrators listed in the Catalog of Navy Training Courses share the goal of providing "the necessary knowledge and skills to

perform advanced level networking system management, administration, and maintenance support," or similar, in their descriptions [139]. As such, sailors responsible for managing, administering, and maintaining the network do not necessarily have the familiarity of being end-users trained to *operate* all the services and software on the network suite's servers [140]. For example, expecting Information Systems Technicians to be proficient in operating much of the specialized software (e.g., medical records software) is unreasonable. The chat server is one of dozens of applications that the administrators are responsible to provide to users. As another specialized software product, the Global Command and Control System – Maritime (GCCS-M) POR requires sailors to have specific training based on the nature of their interaction with the system (system administrator, operator, and watch officer), or the training would not be funded. The requirement for a distributed chat architecture made it into afloat tactical networks but without the metastructuring to engender its use for providing a resilient medium for the Navy's C2 communications.

Culture

P20, an aircraft carrier (CVN) Communications Officer (COMMO), suggests that the root of why Information Professional Officers do not use afloat chat servers combines a risk-averse culture and their inexperience:

> The poor performance history of CANES and senior personnel's risk averse culture have prohibited creative problem solving or the introduction of new ideas. Why haven't more IW [(Information Warfare)] or IP [(Information Professional)] officers advocated to use the local chat server and federate it with the Fleet's directed chat server? Simple. The current talent pool for Information Professional DIVOs [(division officers: the entry-level job for a ship's-company officer)] and CSIOs is plagued with LATXFRS [(lateral transfers: how officers change career fields)] that are still figuring out how to be DIVOs or learning what IP is let alone the systems under their charge. (P20)

Although this sentiment on risk aversion and inexperience is unique among participants, the survey data do not contain any reports of in-use distributed chat architectures to contradict it, even among participants that claim knowledge of the capability. Further, P20's sentiment is consistent with literature on how structural uncertainty (caused here by inexperience) causes undue conservatism because of the possibility for unknown, undesirable outcomes [141]. With greater experience, like P20 demonstrates having, comes greater understanding of the

communications and network architectures, which reduces structural uncertainty and the perceived risk associated with innovating to improve resilience. Because the inexperienced officers described by P20 lack time in the fleet, any intervention to improve chat use must consider administrators' and users' limited time to learn, configure, and teach others about using chat more resiliently.

3.2.3 On Those Who Use Afloat Chat Servers

Although they comprise a minority, understanding how our respondents who use afloat chat servers do so is informative to expanding their use. We list these details in Table 3.2 on how they use afloat chat servers We summarize the following results from the data:

- On the ships where sailors use their chat server most frequently, the survey participants personally did not use it.
- Each of the 5 participants whose ships did not bridge their chat servers claims either to not know of the server's channel bridging capability or that they did not devote the necessary time to figure it out.
- P6 reports the only bridged configuration but describes using the chat server for coordinating troubleshooting efforts among maintainers on the ship, a purpose that does not benefit from a distributed chat architecture but does benefit from using a local server.
- Except for P2, P15, and P50, participants listed in Table 3.2 were on their ships before the chat software stopped supporting the chat federation capability (described in Section 3.2.1), so they could have configured it to operate in a distributed chat architecture.

Table 3.2. Details of Afloat Chat Server Use by Ship (9 of 69 Respondents)

P	Ship Class	Use by Sailors on the Ship	Use by Survey Participant	Purpose	Users Connect From	Bridged	Perceived Effectiveness
P15	DDG	Routinely	Did not use	Operational	On my ship	Unknown	Moderate
P50	LSD	Routinely	Did not use	Operational	On my ship	No	Moderate
P2	DDG	Routinely	Did not use	Unknown	Unknown	Unknown	Effective
P21	CVN	Occasionally	Did not use	Operational	On and off my ship	Unknown	Moderate
P64	CG	Occasionally	Just a few times	Operational & Administrative	On my ship	No	Moderate
P6	CVN	Occasionally	Occasionally	Administrative	On my ship	Yes	Marginal
P43	CVN	Just a few times	Did not use	Administrative	On my ship	No	Ineffective
P60	CVN	Just a few times	Routinely	Administrative & Social	On my ship	No	Ineffective
P1	CVN	Just a few times	Just a few times	Proof of concept testing	On my ship	No	Moderate

Note: Cited Ship Classes: Guided missile cruiser (CG); Aircraft carrier (CVN); Guided missile destroyer (DDG); Dock landing ship (LSD)
Gray shading indicates that the bridging capability was unavailable for these participants.

These nine participants' use of their afloat chat servers and survey responses reveal a lack of awareness of the afloat chat servers' capabilities and benefits. These findings are consistent with the subset of participants whose ships do not use their chat servers.

3.2.4 Discussion on and Recommendations for Chat

To help overcome the primary two reasons survey respondents were not providing their afloat chat servers for use in a distributed chat architecture (unaware of capabilities and benefits, as listed in Table 3.1), empowering the Information Professional Officers to mediate the technology use could help the Navy achieve its resilience goal. Devoting the time and resources to this task is a matter of prioritization. For example, receiving training on distributed chat architectures, which takes time and resources, could empower Information Professional Officers to create guidance, like standard operating procedures, to inform sailors on their ships of the chat channel bridging capability, the resilience benefits of a distributed chat architecture, and how to connect. This is consistent with Orlikowski et al.'s findings, that "mediators create policies, procedures, guidelines, templates, access mechanisms, applications, and physical configurations" [128, p. 437] to influence technology use. These operational aspects differ from the "system management, administration, and maintenance" skills that sailors gain in the official training courses [139], [140]. Such technology-use mediation increases information transparency as mediators provide users with "understandings, images, concepts, knowledge, and heuristics about their specific technology" [128, p. 437]. The ambiguity-conflict model suggests these efforts, in turn, will decrease ambiguity and microimplementation diversity to reduce the policy-implementation gap [9]. Until then, a lack of coupling between the IT adoption policy and its implementation causes the ambiguity to be too great for microimplementation autonomy to achieve the intended implementation (see Figure 2.1).

As a counterpoint, consider that a lack of information transparency or capability requirements has not always hindered sailors' innovation attempts. For example, in 2011, before the Navy used dynamic multiplexing in its afloat IP connectivity, sailors cross-connected SATCOM links based on typical throughput demands to increase the available throughput for the ship [142]. Sailors were watching unused transport capacity on one SATCOM link waste away while suffering through congested throughput on another SATCOM link, so innovating to overcome the issue and improve their communications resilience became a

priority for them. In contrast, with chat, no perceivable problem arises until the ship loses connectivity with the shore and then, the focus is on restoring the ship's connectivity, not analyzing how to configure an application to be more resilient. Without a distributed chat architecture, the loss of connectivity with shore causes a loss of ability to convey C2 communications over chat, as contrasted in Figures 3.3 and 3.4. Resilience is invisible yet critical. Less motivation exists for fixing something that does not appear broken [109].

In the Apple ERP example (described in Section 3.1), the implementation team was too risk-averse to upend current processes for those that promised to be better. The new ERP software, whose implementation was a priority for the CEO, had too much structural uncertainty. The CEO needed to intervene through restructuring to bring about the greater good for the organization. Overcoming junior Information Professional Officers' inexperience to prevail over their risk aversion may similarly need a CEO-level champion, like a fleet commander, who can reframe the chat use case for empowering them to mediate the afloat chat servers' use. A person in this authoritative position can recommend or direct using distributed chat architectures and provide requirements to PORs for system design improvements and explanatory guidance to afloat C4 leadership.

The acquisition system in place when the Navy acquired its afloat chat system differs from today's. The Defense Acquisition Management System in place when the Navy acquired this software yields a gap once the acquired chat capability reached the Operations and Sustainment phase [143]. The requirement owner (i.e., sponsor) did not provide institutional support for the appropriate technology-use mediation so its microimplementation differed from its intended implementation. In today's terms, the Software Acquisition Pathway in the Defense Acquisition System's Adaptive Acquisition Framework prioritizes continuous interaction between sponsors, program managers, and users. Specifically, it directs the sponsor to assess annually "whether the mission improvements or efficiencies realized from the delivered software capabilities are timely and worth the investment" [144, p. 18]. If followed, this feedback from end users to the sponsor and from the sponsor to the POR should prompt the technology-use mediation necessary to expand the software's use [128]. In turn, it would reduce the policy-implementation gap by bringing the realized implementation closer to the intended implementation.

Quantifying this policy-implementation gap with the model in Equation 1.2, according to the procedure provided in Section 1.1, provides a similar perspective as the qualitative analysis. If we model the policy hierarchically as (1) use afloat chat servers (2) in distributed chat architectures (3) for routine fleet operations, then $M = 3$. Let these implementation items relate to the policy requirements: (1) afloat clients connect to their ship's afloat chat server, (2) administrators federate chat servers and bridge channels, and (3) this is the normal mode of operation. The calculation of $cip(I, P)$ can yield 0, $\frac{1}{3}$, $\frac{2}{3}$, or 1 for each survey response, providing a summary representation of policy implementation. Mapping the value earned by each survey response (compare with Table 3.2) yields the following $cip(I, P)$ values [(frequency) value]: (60) 0s, (8) $\frac{1}{3}$s, (1) $\frac{2}{3}$s, and (0) 1s. Aggregating these values by averaging yields $cip_A(I, P) = 0.0483$ on the 0-to-1 ratio scale, with 0 representing no policy implementation and 1 representing full implementation.

Inherent in our recommendation for empowering the C4 leadership to be technology-use mediators is the assumption that C4 leadership will provide and use this capability if they know that afloat chat servers exist and that they can achieve a distributed chat architecture (Figure 3.3) that is more resilient than the prevailing configuration (Figure 3.4). This assumption is a limitation of our recommendation. The only indication in the survey data that a ship might use its chat server in a distributed chat architecture is P65's claim: "I would have ensured the entire strike group would have pointed their chat clients to the server so that it could have been utilized even if cut off from shore." If a fleet commander believes operating chat in a distributed chat architecture is a good idea for C2 communications resilience, it is more likely to happen, like the CEO of a company directing the use of an ERP tool.

To overcome this limitation, we recommend a sociotechnical approach that combines the awareness effort with system design improvements [83], [108], [145]. Such improvements in a Navy chat system would automatically federate servers and prompt users to use a distributed chat architecture. We describe additional system design improvements in Supplemental 2. As the survey results reveal, having capable technology is only part of the solution because users might not use the intended resilience-providing capabilities. But if the technology were more capable and supportive of the policy for resilience, less ambiguity would exist about the intended configuration.

3.3 Perceptions of Communications Shift Procedures

U.S. Navy ships depend on having resilient IP connectivity with off-ship resources. While shifting these services from one area of responsibility (AOR) to another, the Navy desires to minimize any negative effects of the shift. However, ships' network administrators sometimes use suboptimal procedures for shifting their SATCOM and IP network services, leading to unused IP paths. To discover why they do not always optimize bandwidth availability in such shifts, we surveyed four sailors on ships conducting cross-AOR shifts (survey instrument shown in Appendix B). To our surprise, more pressing challenges hinder ships during their shifts. Instead, simple completion becomes the goal, reducing or eliminating opportunity for any optimization effort of the communications shift. Such a finding also prevents us from using a model, like Equation 1.2 offers, to quantify the policy-implementation gap.

3.3.1 Communications Shift Procedures

We focus this portion of our research on ships' communications-shifting procedures during CHOPs involving changing NOCs. In summary, the traditional procedure used with the older architecture involves shifting all SATCOM first, followed by the network enclaves. Using the older procedure in the newer WAN multiplexing system architecture (shown in Figure 3.1) enables ships to use previously unusable portions of bandwidth [116]. Optimizing the procedure involves shifting one or more networks to the gaining NOC first, followed by SATCOM, and finishing with one or more network shifts. We describe the architecture and procedural details in Supplemental 2.

3.3.2 The Surveys

Because participation in our survey must coincide with a cross-AOR CHOP during a ship's deployment, the population of available participants is very limited. Ships must have at least two SATCOM missions (not counting extremely high frequency (EHF) Time-division multiple access (TDMA) Interface Processor (TIP) connections) to shift when conducting a CHOP to another AOR. To find ships eligible for our survey, we reviewed fleet schedules on the SIPRNet Collaboration at Sea (CaS) servers about weekly for about a year. For each eligible ship, we queried the Regional SATCOM Support Center (RSSC) wideband and Naval Network Warfare Command (NNWC) Commercial Broadband Satellite

Program (CBSP) SATCOM assignments databases on the SIPRNet Intelink and U.S. Fleet Forces SharePoint servers, respectively. We solicited ships through their TYCOM. Although having the TYCOM staff solicit ships gave credence to our survey, to avoid coercion, we clarified to sailors that their participation in the survey was voluntary and they would suffer no penalty or loss of benefits for not completing it. In our solicitation, we requested that the communications officer or someone with similar knowledge and authority complete the survey.

We solicited all 12 ships that met the survey eligibility criteria between July 2023 and June 2024, getting four responses. We invited ships to participate in the survey for each CHOP during their deployment for which they were eligible; however, none volunteered. No response (provided in Supplemental 3) met the eligibility criteria and was complete. Survey participants' answers to the qualitative portion of the survey, including their elaborations on their experiences in the open-ended survey questions, reveal aspects of the cross-AOR shift process that we may have otherwise overlooked. For example, they identify who directed their communications-shift procedures and their satisfaction level with the process.

We highlight some characteristics of the received surveys here. One ship (S1) lacked enough SATCOM missions to be eligible for the survey but submitted a complete survey anyway. Two responses are incomplete because respondents answered question 4 partially, either omitting a subset (S2) of or all (S4) the SATCOM missions or the IP network enclaves (S3). For example, S2 identifies a SATCOM mission in question 1 that shifts to a different configuration in question 2 but omits that shift in question 4. We suspect that S2 reported only the SATCOM shifts that involved changing satellites because the shifts that did not involve changing satellites would have been mostly transparent to the ship while the losing NOC re-patched the connection to the gaining NOC. Not knowing the order or timing of the steps prevented us from measuring the bandwidth lost in the shift. Three ships (S2, S3, and S4) shifted their SATCOM missions and IP network enclaves in the order directed by a higher authority, either a group commander (S2 and S3) or Joint Fleet Telecommunications Watch Officer (JFTOC) Watch Officer (co-located with a NOC) (S4). Future research can investigate the group commanders' level of understanding of how the afloat tactical network architecture enables staying online when ships have multiple SATCOM missions simultaneously. For S1 and S4, they claim that the order in which they executed their shifts

was the same as it always is (Question 5, Appendix B), indicative of a status quo bias, as described in Section 3.1.

3.3.3 Communications Shift Experiences

On average, survey respondents are "neither satisfied nor dissatisfied" with their experience during the shift. However, on average, they see "a great extent" of room for improving the quality of the SATCOM and IP network communications shift process. Because our survey respondents are not randomly selected, we must consider the possibility for a self-selection bias.

Despite operating a WAN system for decades for spanning IP networks between ships and shore facilities, the Navy still struggles to make SATCOM and IP network communications shifts routine. One survey respondent said:

> For something that happens consistently to all ships as they maneuver throughout the ocean, I was expecting a routine process to follow that would ensure services shifted in a timely and effective manner. Instead[,] it behaves similar to any other troubleshooting event, where the ship struggles to connect with shore entities to isolate and resolve problems. (S2)

Respondents noted two categories for possible system improvements, satellite access authorization (SAA) administration and equipment or system-specific issues. Three respondents (S1, S3, and S4) experienced unplanned outages during their communications shifts. At the macroimplementation level, these experiences involve low ambiguity and low conflict (execute the procedures to accomplish the mission) so implementation should occur simply when all the necessary resources are available (see Figure 2.1) [9]. However, at the microimplementation level, these experiences demonstrate high ambiguity in the execution details and high conflict in contention for limited satellite resources. As such, implementation occurs when sufficient understanding forms between a "coalition" of personnel aboard the ship and at the various supporting stations ashore.

To get satellite access, ships submit satellite access requests (SARs) at least 45–60 days (varies according to region) before they can connect. RSSCs evaluate and rank all SARs against national priorities, as brokers of a limited resource. Upon approval, the servicing

RSSC issues a corresponding satellite access authorization (SAA) that assigns the SATCOM mission parameters, including the satellite resources, start and end times, ground station (called a Standardized Tactical Entry Point (STEP) site) terminal configuration, and the requesting ship's SATCOM terminal configuration.

S1 and S4 mentioned how a lack of an SAA before they need SATCOM resources caused trouble. For S1, receiving the SAA only after its previous SAA ended caused a 78-hour wideband SATCOM outage and inhibited planning the shift with the gaining NOC. The survey data do not reveal whether S1 submitted its SAR on time, schedule changes, or other factors outside its control caused this gap in SATCOM coverage. Short-notice schedule changes delayed S4: "While departing from the pier S/F [(ship's force)] experienced delays in getting DoD WB [(wideband)] approved due to a surge deployment and could not identify a sponsoring component."

For S3, the STEP site cut the ship's SATCOM access before the SAA ended and reallocated the SATCOM resources to another unit: "The [SATCOM] mission was secured [terminated] prematurely with no explanation" (S3). Because another unit was using those resources (apparently double booked), S3 could not get them back and suffered a 28.8-hour outage, lasting until its next SAA started and S3 could establish IP connectivity through it.

Another factor causing problems for S3 was its older systems: "All [people at the] comms stations we used on deployment . . . were unfamiliar with our ship's capabilities and limitations, despite conducting planning teleconferences where our concerns were specifically laid out." Specifically, during S3's wideband outage, S3 needed to shift its protected-band SATCOM. (*Wideband* and *protected-band* are military SATCOM terms representing differences in frequency and waveform; both are suitable media for ships' IP WAN connectivity.) However, the gaining NOC was unprepared to connect with S3's older protected-band terminal, which enabled a 118.4-hour outage. The wideband and protected-band outages overlapped, causing a 24.2-hour complete IP service outage for S3 while deployed.

If S3 and the gaining NOC discussed S3's equipment configuration, how could the NOC personnel be unprepared to connect? Those who have served at communications stations would be quick to suggest that the personnel involved in the planning teleconference simply forgot to pass the information to the team on watch [146], [147]; however, this seems unreasonable to be the root cause of an outage lasting nearly five days. Neither can we call

this NOC consistently substandard because S2 praised this same NOC for its troubleshooting support. The survey data do not identify a root cause for this outage.

3.3.4 Respondents' Ideas for Improvement

S1's and S4's comments suggest that the complications in getting SAAs indicates that the system insufficiently supports deployed ships, thus increasing ambiguity (not knowing whether a shift will be adequately supported) and conflict (as an unplanned outage incurs additional risk to the ship's mission) at the microimplementation level. S1 recommends ships and NOCs coordinate better and earlier; however, even when ships discuss communications-shift plans with the involved NOCs, problems can persist, like S3's experience. S1 also recommends that NOCs should continue to track a ship's shift progress through completion: "It seems like the process should be something that is tracked until the unit CHOPing has switched over successfully." Such "tracking" would help engender the "coalition" necessary for policy implementation in a high-ambiguity, high-conflict scenario [9].

Of the multiple factors affecting a ship's shift schedule, S3 suggests that ships have greater input. A communications shift typically occurs within 24 hours of a CHOP and ships (or strike groups) have some control over the timing of their shifts in their SARs and IP Service Requests. If a ship believes the most convenient time for a shift is at 1500Z, then it need not request a service end or start time at 0001Z just because it marks the beginning or end of its assigned CHOP date. Ships have greater flexibility when they coordinate with the Wideband Satellite Operations Center (WSOC) and JFTOC during their shift and could extend their satellite access beyond the SAA end time or start it earlier than the SAA start time. Such coordination strengthens a coalition at the microimplementation level. Ships have even greater flexibility in shifting their network enclaves (including NIPRNet and SIPRNet) and, if prepared, could shift them without NOC support [116]; though, this may increase conflict between the ship and NOC, harming the coalition and potentially causing additional problems.

Both S2 and S4 describe a practice of shifting their SATCOM and IP network services before getting underway on deployment from their home ports. This procedure puts the potential outage before the deployment starts and enables easier access to the ship for technical help from shore stations. Also, it is consistent with the older Navy WAN architecture's

procedures when ships needed to completely isolate from one NOC before connecting to the next, as described in Section 3.3.1. However, with the newer WAN architecture, no outage requirement exists for shifts. The survey data do not indicate whether procedure S2 and S4 describe is an outcome of the status quo bias (i.e., the way they have always done it) or a risk mitigation measure (i.e., to avoid a potential outage while deployed).

These experiences help us answer the initial research question for this study about why some sailors are not using the resilience-improving capabilities of deployed IT. The ambiguity-conflict model helps explain that system complexity, ambiguity, and conflict at the microimplementation level all contribute to the Navy's inability to ensure smooth communications shifts routinely [9]. The model also supports sailors' ideas from the survey data for improving communications shifts during CHOPs.

3.4 Discussion and Conclusion

How can the level of adoption of available IT capabilities contribute to gaps between policy and practice? We explore the Navy's adoption of IT systems that can support resilient command and control (C2) communications: the Navy's afloat chat system and WAN architecture. We find that these systems were not fully "adopted," that is, completing their acquisition, implementation, and initial user training [13]. Because their implementation and user training were incomplete, their operation in the continuance phase of the information system life cycle revealed a policy-implementation gap.

The ambiguity-conflict model [9] and the concept of metastructuring through technology-use mediation [128] help explain surveyed sailors' underuse of resilience-providing capabilities and form the basis for our intervention recommendations. For chat, a decoupling of the Navy's goals for the system and the system's implementation has caused a policy-implementation gap between them. For the communications shifts, high ambiguity and high conflict at the microimplementation level enables success only when sufficient understanding forms between a coalition of personnel aboard the ship and at the various supporting stations ashore. In both cases, high ambiguity has the greatest impact on the diverse microimplementations that widen the policy-implementation gap.

In reducing ambiguity for each system's implementation, we recommend a sociotechnical approach, that is, an approach that views "society and technology together as one coherent

system" [104, p. 1] because focusing only on the social or only on the technical will not overcome the nonuse that we observe. Such an approach requires the U.S. Navy's leadership's intervention to use the full capabilities of applications to maximize their resilience. We recommend the Navy continue executing the multiple tasks described in Supplemental 2 to improve the systems' design and resilience of its C2 communications. Realizing the operation of distributed chat architectures requires both a technical configuration change and mediating technology use for the resilient chat capability. Survey data reveal that the most significant hurdles to using afloat chat servers include respondents' awareness of their existence, capabilities, and benefits and how best to use their capabilities, indicating insufficient metastructuring.

One may argue that people sometimes do not adopt a technology because they believe that its disadvantages outweigh its benefits such that the perceived benefits are not worth the effort to adopt it [for examples, see 108, pp. 149–150, 244–245, and 278]. Recognizing a technology's worth can be difficult when the technology's users do not know about its capabilities and benefits and such benefits help only in contingencies, like a communications loss or degradation [109]. However, our survey data do not contain an indication that anyone sees a disadvantage of operating a distributed chat architecture or using improved communications shift procedures.

One may suspect that other capabilities of deployed systems also go unused because they could be unfamiliar to sailors. Promisingly, the relatively new Software Acquisition Pathway in the Defense Acquisition System's Adaptive Acquisition Framework directs an annual assessment mechanism that forms the very structure to reveal gaps like this to sponsors [144]. Future research can explore the effectiveness of this newer acquisition method. Technology-use mediation to raise awareness will involve providing greater understanding into why ships have afloat chat servers and how to operate them in a distributed chat architecture by having a fleet commander-level champion promote their use and circulate guidance. This circular relationship of receiving feedback and implementing it throughout the fleet comprise an effective metastructuring approach [128]. Whether the Adaptive Acquisition Framework's annual assessment mechanism will yield tighter links between development and use can be the subject of future research.

In seeing the lack of realized resilience, it may seem easier to decide to adopt a new technology than fix and maintain the old technology, which still might not achieve the intended resilience [148]. U.S. Navy engineers are developing Communications as a Service (CaaS) to enable an *anything over anything* capability and are integrating new chat software that uses CaaS for its connectivity between afloat networks. It enables encoding IP traffic for conveyance over paths not traditionally recognized as IP paths. For example, one CaaS server can convey IP traffic to another CaaS server in J-series messages over a tactical data link (TDL). The capability promises to be more resilient than previous chat capabilities because of its ability to use many communications paths, not just native IP paths. Will this become yet another unused capability like the distributed chat architectures that the fleet first used in Trident Warrior 2005 [111]? If so, then the policy-implementation gaps revealed in this chapter will likely persist with the new system. What technology-use mediation and system design actions will the Navy empower its sailors to take to reduce these gaps and help ensure its success? Learning from the lessons observed in surveyed sailors' (non)use of afloat chat servers, the Navy should ensure that sailors (C4 leadership and end users) are aware of the new chat capability, its benefits, and how best to use it to apply those benefits to bolstering the resilience of C2 communications. We must therefore understand the whole-system context for adopting each new system.

As the U.S. Navy emphasizes command and control (C2) communications resilience in distributed maritime operations and communications-degraded or denied environments (CD2Es), understanding current systems' use can reveal better ways to use them. Further, applying these lessons observed to new systems can help keep new systems from the same fate (of their resilience-providing capabilities being underused) and improve warfighter effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4:
Using a Commonly Accepted Framework to Protect Against Privacy Inferences

An adapted version this chapter appears as "Mitigating Inference Risks with the NIST Privacy Framework" by C. B. Landis and J. A. Kroll in the *Proceedings on Privacy Enhancing Technologies (PoPETs)*, vol. 2024, no. 1, pp. 217–231, 2024. Available: https://doi.org/10.56553/popets-2024-0013 [45].

Organizations face substantial challenges in practical privacy risk management. For example, for over a century, the U.S. Census Bureau has faced privacy risks in its lawful collection of personal data on U.S. citizens [149]. A practice the Census Bureau follows is to not release raw data, only products, like statistics. While privacy-preserving methods have evolved over time, so have adversarial capabilities to infer identities and attributes of ostensibly deidentified data [150]. Formerly promoted deidentification methods for statistical databases [151], [152] are no longer sufficient [153], [154]. Although privacy-enhancing technologies (PETs) can help, many organizations struggle in their use of risk controls [155].

Organizations face increasing levels of risk when handling sensitive data. In this context, we explore the research question, "How can standards or frameworks guide or hinder organizations' policy implementation?" The National Institute of Standards and Technology (NIST) Privacy Framework [43] is an important privacy-focused risk management tool in the U.S. While some organizations have benefitted from using it [44], NIST offers a caution against using the framework as a compliance or assessment checklist. Risk governance within organizations is often focused on legal compliance [33], [156], which provides limited privacy protection. Despite the many legally mandated data protection practices, including those of the General Data Protection Regulation (GDPR) in Europe, comparable laws in the U.S. (e.g., the California Consumer Privacy Act (CCPA) and Virginia Consumer Data Protection Act (VCDPA)), and other jurisdictions (e.g., signatories to the Council of Europe's Convention 108), some decisions, like placing a privacy team in an advisory role rather than an embedded role, can hinder privacy efforts [33]. They can impede an organization's risk professionals from foreseeing compliance problems. The NIST Privacy

Framework can help organizations in addressing this gap. We apply Matland's ambiguity-conflict model of policy implementation to explain how this gap can occur [9].

The framework fails to capture key ideas about specific privacy risks from the privacy research community. It considers mostly the risks borne of unauthorized information flows and cybersecurity failures while under-addressing data *inferences*. This gap limits the positive outcomes of practical privacy risk management. For example, Target inferred that a teen was pregnant by her ordinarily non-baby-related purchases and sent coupons to her for baby products, which antagonized the teen's father who was unaware of his daughter's pregnancy [8]. News of the incident spread quickly and caused bad press for Target. One may argue that this shortfall is like that in steganalysis, in that steganographers will always be a step ahead [157]. Although a similar arms race may exist between deidentification and reidentification techniques, several compelling inference incidents occurred years before NIST publishing its Privacy Framework, so the framework should have addressed inferences.

Much privacy research has focused on the risk of *reidentification* of anonymized data, often provoked by the popular notion of redacting personally identifiable information (PII) for deidentification. However, over a decade ago, researchers established that protecting privacy by protecting only information in PII categories is ineffective [158], [159]. It is insufficient for preventing the reidentification of an individual in deidentified data, as described throughout this chapter. For example, Cohen demonstrates the insufficiency of k-anonymity and other quasi-identifier (QI)-based anonymization techniques [153] on the EdX data set [160]. The gap between legal requirements and technical implementations enables workarounds, both unintentional and intentional, to the deidentification process [17], [153], [161]. As such, the community has largely refocused its efforts around provable guarantees and quantifiable metrics, either of formal indistinguishability properties like differential privacy [162] or of system-level information-hiding properties [163], [164].

The NIST Privacy Framework [43] describes itself as a comprehensive approach to an organization-wide privacy program. It aspires to enable dialogue among executives, managers, and practitioners to organize, assess, plan, and execute a privacy program in any organization, customized to that organization's needs. Frameworks like the NIST Privacy Framework should help organizations balance the conflict between increasing the utility of data by releasing them and protecting the privacy of people by safeguarding them. The

framework authors assert that cybersecurity can exist without data privacy safeguards but that protecting privacy is impossible without effective cybersecurity [43]. It maps categorized risks to controls [165], which primarily focus on aspects applicable to cybersecurity. Cybersecurity and the framework's mapped controls still fall short of sufficiently managing the risk of attackers inferring attributes [166], for example, as in mosaic theory [167]. No statute or regulation requires framework adoption.

We analyzed how well the NIST Privacy Framework identifies and controls the risks of sensitive data inferences by applying it to four inference incidents. The framework does not guide organizations' attention to enough of the various inferences risks that they should mitigate, depends too much on an obsolete practice (i.e., protecting privacy by protecting PII categories), and has limited guidance for translating policies into technical implementation. We establish this taxonomy to describe two forms of inference risks:

Reidentification Inferences: Associating the original individuals with data that have been anonymized, including by removing PII [159]. For example, researchers reidentified individual users in a Netflix-released data set, even though Netflix had deidentified the data before its release [168].

Operational Inferences: Inferring sensitive attributes about people and organizations, like the Target teen-pregnancy example [8]. Even with perfect disclosure controls to prevent reidentification inferences, operational inferences unrelated to PII may still be possible, depending on the threat model [169], [170].

We introduce privacy inferences in Section 4.1 and the NIST Privacy Framework in Section 4.2. Our primary contribution in Section 4.3 is to validate the framework's capacity to identify and mitigate inference risks by applying it to four incidents. In Section 4.4, we propose recommendations for how organizations could have disrupted these incidents' structural and systemic causes. We conclude in Section 4.5.

4.1 Privacy Inferences

For decades, many believed that managing one's privacy consisted of a relatively simple framework of limiting the sharing of sensitive information and being informed of and consenting to any other entity's use of sensitive information. Advancements in information technology (IT) and using *quasi-identifiers (QIs)*, however, have proliferated inference-

generating capabilities beyond the control of any individual [171]. "Quasi-identifiers are sets of attributes that are potentially available to an attacker from other sources, combinations of which may uniquely distinguish an individual within the dataset" [160, p. 1469], for example, the list of courses in which a student is enrolled. Privacy, then, has become too complex for an individual to handle [172]. Such capabilities have even destroyed the foundation of formally trusted deidentification techniques [159], including k-anonymity. k-anonymity means that "the information for each person contained in the [deidentified data set] release cannot be distinguished from at least $k-1$ individuals whose information also appears in the release [173].

Today, organizations desiring to protect their members' and constituents' privacy should develop a comprehensive privacy-risk management program, and the NIST Privacy Framework can help. As privacy protections have evolved so have people's definitions of what privacy means to them, which complicates privacy protections [42], [174], [175]:

> An effective privacy risk model must capture the potential cost of problems to individuals, but it may be difficult to do with any consistency because there may be a significant divergence in the way individuals experience problems; that holds true especially for embarrassment or other psychologically-based problems. Assessing the impact on individuals is an area that needs further research. . . . Context, therefore, is a key characteristic in determining the likelihood of a data action becoming problematic. [42, p. 22–23]

4.1.1 Inference Incidents

Past incidents demonstrate how people and organizations have been harmed by inferences. Studying them offers opportunities for learning how to mitigate future incidents. Our taxonomy categorizes inference incidents as either reidentification or operational incidents.

Reidentification Inferences: MIT and Harvard EdX Data

In 2014, the Massachusetts Institute of Technology (MIT) and Harvard published data of student online course performance on their EdX platform. They anonymized the data using QI-based deidentification to achieve k-anonymity, consistent with the requirements of Family Educational Rights and Privacy Act (FERPA) [176]. Nevertheless, the data remained vulnerable to reidentification attacks. For example, a *downcoding attack* reverses a QI-based

deidentification's hierarchical generalization by logical deduction, thus revealing portions of the deidentified data. Another example is a *predicate singling-out* attack, combining the notion of a "predicate," which is synonymous with QI, and "singling out," which is synonymous with a reidentification inference [153]. Thus, a predicate singling-out (PSO) attack involves inferring a person's identity from QIs in deidentified data. In *highly dimensional* data sets, each record has at least dozens of attributes, like in the Netflix data set, social media friendship graphs, and genetic data [168], [177]. The greater the number of dimensions, the greater the opportunity to form QIs and find unique records. Cohen [160] could reidentify anonymized records by executing downcoding and PSO attacks on the highly dimensional data set.

Cavoukian and Castro argue that because demonstrations of these types of attacks appear primarily in computer science venues, such attacks require computer science expertise and rely on insufficient deidentification techniques [178]. Many refute this claim [158], [159], [177], [179]. Cohen emphasizes and shows how people without specialty education, experience, or tools could reidentify students' records. A "prospective employer, a casual acquaintance, and an EdX classmate" [160, p. 1478] are among the potential adversaries that could figure out whether a person failed any EdX courses and use that information harmfully. The harm can include denial of job opportunities and public embarrassment.

Cohen speculated that the EdX experts deidentified the data set on one attribute after another, followed by selective row deletion. They claimed to have achieved 5-anonymity across 17 overlapping QIs by removing records associated with 20% (121,160) of 597,692 students; however, the union of these 17 QIs yielded 7.1% student uniqueness (1-anonymity) in the data and 15.3% of students with less than 5-anonymity. The "disjointed," rather than atomic, deidentification process degraded the initially established k-anonymity [160]. Furthermore, the widespread technique of using QI-based deidentification on highly dimensional data sets can preserve enough information to enable reidentification whereas differential privacy techniques can overcome this [153]. QI-based deidentification relies on "unstated assumptions on the data distribution" [160, p. 1480] and likewise, is insufficient to mitigate reidentification, even when used in FERPA-compliant anonymization. Dwork and Roth comprehensively define *differential privacy* and summarize it as arriving at the same conclusion about a data set regardless of whether a record is in the set [162]; we expand this definition in Section 4.1.4. If FERPA's authors had included reidentification inference

threats or if the EdX data set deidentification had used an atomic rather than a disjointed process, reidentification would have been much more difficult [180].

Reidentification Inferences: NYC Taxi Data

In response to a Freedom-of-Information-Law request, the New York City (NYC) Taxi and Limousine Commission (TLC) released its taxi data from 2013 [181]. It anonymized the data by hashing some sensitive fields, including driver license and medallion numbers, which can reveal drivers' annual income levels. A dictionary attack (creating a table of hashes based on the format and range of possible values [182, p. 269]) quickly resolved the hashes to their original, sensitive values. For inferences, though, analysts can reidentify passengers and infer patterns of life, whether favorable, benign, or unbecoming, by combining the remaining data (pick-up and drop-off times and locations, itemized payment amounts, and driver identification) with information that is public or otherwise known. Public information in this deidentification case includes news media, municipal public property records, and social media accounts. Published examples of privacy inference reidentification include stalking celebrities, as shown in Figure 4.1, and reidentifying a man who visited multiple strip clubs whose life could be ruined by someone who discovered these travels [183]. The harm for celebrities can include associations with businesses and residents that the celebrities would have preferred to keep private.

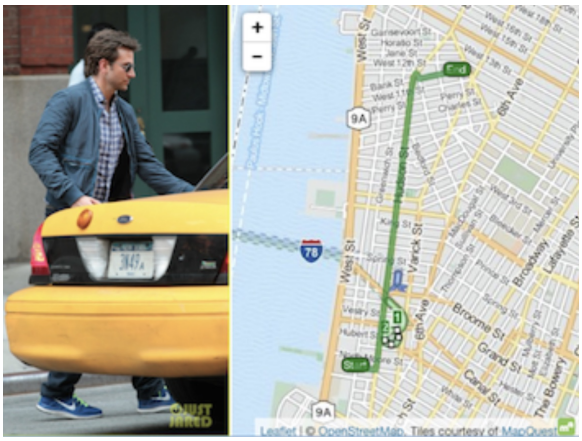


Figure 4.1. A Public Photo of Bradley Cooper Getting into a NYC Taxi Enabled Inference Reidentification of His Corresponding Taxi Route. Source: [183].

Although the shortfall in deidentifying the taxi data such that passengers can be reidentified shows a disparity between complying with the law and achieving the anonymity goal [17], [161], the opportunity for inferences poses a greater risk to people and liability to the Taxi Commission. The Taxi Commission ostensibly complied with the law when deidentifying the data before its release, yet passengers, who had no personally identifiable attributes in the data set, can be reidentified. Tockar demonstrated how differential privacy techniques could have helped prevent inference-based reidentification [183]. However, Douriez et al. demonstrated that differential privacy is insufficient after transforming the taxi data set into a moving object database of trajectories [181]. Similar reidentification opportunities exist today [184], [185], [186].

Operational Inferences: Strava Heat Map

The Strava fitness app collects geolocation tracking data from its users and aggregates them to generate a global heat map of workout locations, the brightness directly proportionate to the popularity [187], enabling operational inferences. Only days passed between Strava publishing its heat map in 2018 before people noticed workout paths in the Afghanistan wilderness, shown in Figure 4.2. An easy inference identifies them as coalition forward operating bases, inhabited by deployed military personnel. The heat map also revealed a covert CIA site near Camp Lemonnier in Djibouti, an inference which could be fatal to the lives of the agents.

Figure 4.2. Strava Workout Tracking in Helmand Province, Afghanistan, Determined to Geolocate a Coalition Military Base by Operational Inference. Source: [188].

Strava deidentified the data for its heat map by aggregation. Nevertheless, the aggregated data violated the privacy of groups of individuals [187]. Strava noted that users could opt-out of sharing and avoid revealing sensitive locations [187]. In other words, individuals' seemingly innocuous behaviors harmed their organizations' operational security (OPSEC). Still, perhaps the best place to limit the spread of sensitive information is at its source. Many communities use OPSEC principles but, the U.S. Government (USG) and the Department of Defense (DoD) have a detailed OPSEC program policy available publicly [20], [170].

Operational Inferences: Pentagon Pizza Orders

In 1998, the owner of 59 Domino's franchises in the DC area had a "Washington Pizza Index," the number of pizza orders that he correlated with the level of USG activity [189]. The index revealed and enabled operational inferences. Furthermore, based on the area of the order, he could infer whether the pizza was headed for the White House, congressional offices, or the Pentagon. For instance, the most pizzas ordered per day was by people working in the Pentagon during the Persian Gulf War [189]. During a war, one might expect a military's headquarters to be working longer hours than normal. However, extra hours devoted to planning before a major operation may reveal too much, potentially compromising the success of the operation by signaling its imminence to adversaries. Likewise, a corporate

build-up working toward the launch of a new product could provide similar inference opportunities. Similar clues include parking lot fullness and employee shuttle activity levels, and so on can likewise provide opportunities for unwanted inferences.

Lessons Observed

The NYC taxi passenger reidentification incident shows how attackers can combine anonymous data with public information from sources beyond the affected organization's control. Similar patterns of reidentification inference occurred in the Netflix prize [168], AOL search history [190], and Massachusetts Group Insurance Commission [191] data release incidents. Our other reidentification incident, EdX, revealed course performance. The Strava and pizza incidents show how data aggregation can worsen privacy protections. Similar patterns of operational inference occurred from electrical smart meter activity because of an electric company's lack of protecting sensitive data [192] and Cambridge Analytica inferring Facebook users' political affinity (an attribute), causing penalties for Facebook [6], [193].

The themes arising from these incidents in better mitigating reidentification inferences were largely technical and those in better mitigating operational inferences were largely sociotechnical. The following contributed to the inference reidentification of people in deidentified data:

- Keeping too many quasi-identifiers (QIs) (EdX, Netflix, MA Health)
- Inability to translate legal requirements unambiguously into technical implementations (EdX, NYC)
- Disjointed, rather than atomic, deidentification (EdX)
- Lack of understanding deidentification methods (NYC, AOL)

Sociotechnical aspects, including human behavior and configuration decisions were the primary contributors to enabling operational inferences. The harmed organizations (coalition forces, CIA, and DoD) in these operational inference incidents were involuntarily harmed by the behavior of their members (deployed and headquarters forces) and third-party organizations (Strava, Domino's, and the Washington Post). The harmed victim is often not the primary organization involved in these unexpected OPSEC failures. The root of these observations is inadequate consideration of the potential consequences of actions, including:

- Risks of sensitive inferences from personal data were not identified (all operational inferences)
- The by-design end-user defaults concealed risks (all operational inferences)
- Aggregate irregular behavior revealed sensitive information (Strava, Pizza, Smart Meter)

4.1.2 Accounting for Sensitive Inferences in Law

Unlike some notable U.S. privacy and privacy-related laws that do not mention inferences as a threat to privacy, we expect that newer privacy laws will address inferences. Researchers have reported this shortfall in the Health Insurance Portability and Accountability Act (HIPAA) [194], [195], Children's Online Privacy Protection Act (COPPA) [196], [197], and FERPA [176], [180]. The effects of inference attacks as privacy risks are more significant than the way the law currently treats them. Nearly all legal references to inferences that we found relate to reidentification, not operational inferences.

Various researchers propose defining a hierarchy of harms determines the sensitivity of personal data [198], [199], [200], amplifying that some privacy violation harms are intangible [201]. Holder analyzes the GDPR and CCPA, relatively progressive privacy laws, finding that they are too ambiguous or contradictory about inferences [198]. Clifford, Richardson, and Witzleb found growing consensus that lawmakers must address artificial intelligence (AI) and machine learning (ML) technologies from a privacy perspective to protect inferences [199]. Likewise, Feng, Wang, and Chen conducted a similar analysis and proposed new laws in China to account for the privacy of inferences [200].

Wachter and Mittelstadt recommend improving the GDPR to account for protecting against inferences by establishing an individual's "right to reasonable inferences," i.e., a "right on how to be seen" as complementary to a "right to be forgotten" [202, pp. 613–614]. They conclude that inferences are personal data, which requires the "right to know about, rectify, delete, object to, or port" [202, p. 499], except that deletion can come with prohibitions on (re)creating new inferences. They also investigate organizations' facilitation of people exercising their right to their personal data, called subject access requests. If this would reveal an organization's trade secrets or intellectual property, then they recommended that organizations not be obliged to comply [202]. For example, Target revealing specifically

which products lead it to infer that customers may be pregnant would reveal a trade secret because the store that can offer the best deals to an expecting mother will likely gain her business [203], [204]. On the other hand, Edwards and Veale argue that this exception should not be needed because the explanation of an inference-generating algorithm is irrelevant to individuals' rights [205]. Instead, they claim that the rights to deletion and porting are far more important.

The Office of the California Attorney General (AG) came to a similar conclusion about the CCPA [206], asserting that a consumer has the right to know about the inferences that an organization generates [207]. They also agree that when a disclosure would divulge the organization's trade secrets or proprietary methods, organizations are not obliged to fulfill the consumer's request. They also claim that no provision of the then-forthcoming Consumer Privacy Rights Act (CPRA)-revised CCPA would change the published opinion [207].

4.1.3 Inferences for Marketing

Although many business models depend on inferences, including determining creditworthiness, insurance risk, and suitability for matching with an employer, mate, etc., perhaps the most prolific field is marketing [208]. For example, Google [209], Facebook [210], Proctor & Gamble [203], and other companies have had large returns on their targeted advertising. While identifying people's potential shopping interests, based on inferences from their metadata or activities, is not typically a malicious act, whether we should condone this type of inference is controversial [8], [33], [203], [211]. Nevertheless, organizations continue to target their advertising at specific users when they can because it is more effective [210], [211]. For example, Tim Hortons and other fast-food restaurants sometimes track their app users' location at frequent intervals to know when they might be patronizing competitors and to lure them back with special offers [212]. Further, with only the email address of a specific victim user, it is also possible for attackers to infer their selected victim's affiliations and infer their demographics by entangling the victim's advertising identity [213]. Although targeted advertising often involves proprietary methods and data that are not intended to be released beyond an organization's analysts, organizations often share inferences with their contracted third-party analysts, circumventing sharing restrictions [193].

4.1.4 Privacy-Enhancing Techniques

Deidentification is non-trivial [152], [154], [177]. Narayanan, Huey, and Felten offer a "precautionary approach" to safeguarding privacy in data sets, claiming that future reidentification capabilities may be unknowable [179]. Their approach is all about decision-making. The precaution is that organizations should not use any ad hoc deidentification techniques. Instead, they argue for organizations to adopt provable privacy (for which they identify differential privacy as the only option today), consider narrowing the scope of the data to be released, and inform the affected parties of the content of and purpose for releasing the data, sometimes called *accountability notifications*. *Privacy-enhancing cryptography* can be any of several schemes that enable data processing without revealing the data, including differential privacy, homomorphic encryption, and private information retrieval [214], [215]. In turn, we introduce these privacy-enhancing cryptography schemes along with contextual integrity as techniques to reduce the opportunities for inferences.

Differential privacy is one of the most prevalently discussed privacy-enhancing technologies (PETs) today [216]. The formal definition of differential privacy fills a book chapter so we summarize it here analogously [162]. *Differential privacy* provides a mathematical guarantee about an adversary's ability to deduce the values in a database row, that is, the adversary would make the same conclusion about a database regardless of whether a record is in the set. Stated differently, "Differential privacy promises to protect individuals from any *additional* harm that they might face due to their data being in the private database x that they would not have faced had their data not been part of x" [162, p. 20]. The process of applying differential privacy introduces randomness like social science's method of coin-flipping for survey respondents answering sensitive questions about illegal or taboo topics: tails means to answer truthfully; heads means to flip another coin that determines the respondent's yes/no answer. Differential privacy can protect against both reidentification and operational inferences made from differentially private databases, including by neutralizing linking with other data sets and composition of multiple queries. It also enables data owners to quantify an upper bound of their data losses before releasing the data.

Garfinkel, Abowd, and Powazek report on the U.S. Census Bureau's work in implementing differential privacy for the 2020 census, highlighting the challenges experienced [150]. Their three recommendations for furthering the implementation of differential privacy into the Bureau's work are not technical, but relate to organizational communications, bridging

management and technical personnel. Abdu et al. concur, that organizations adopting differential privacy must take a sociotechnical approach to have a chance at meeting success [217]. This parallels many cybersecurity and privacy controls in NIST Special Publication (SP) 800-53, which affect multiple levels of implementing organizations [46].

Homomorphic encryption enables computations on ciphertext data such that the decrypted result matches the result as if the same computations had been completed on the plaintext data [218]. Such a scheme would enable third-parties to process and store sensitive data as ciphertext, thus maintaining data confidentiality. Because the first homomorphic encryption schemes were exponentially slow compared to their plaintext equivalents, researchers have been improving their performance. Further, they have been making the schemes more practical for use beyond research laboratories by, for example, making software libraries available publicly [219].

Private information retrieval inhibits database owners from seeing the queries made by users [215], [220]. In other words, the users benefit from the privacy protection. Some theoretical examples of private information retrieval combine it with differential privacy or other privacy-enhancing cryptography schemes to protect the privacy of multiple parties simultaneously.

Nissenbaum applies contextual integrity (CI) to privacy protection. *Contextual integrity* "is preserved when information flows generated by an action or practice conform to legitimate contextual informational norms; it is violated when they are breached" [221, p. 224]. Analysts identify each flow by the five-tuple {Subject, Sender, Recipient, Information Type, Transmission Principle}. Raw data at the bottom of the flow hierarchy are used by analytic processes at higher levels, generating inferences. From these, analysts can form additional, higher-level inferences or predict instantiation of lower-level raw data. For example, Target inferred that a customer was pregnant by analyzing her shopping history (raw data up to inference) and predicted which products she might consequently purchase while pregnant (inference down to probable raw data) [8].

Martin and Nissenbaum use contextual integrity to explore the important question of privacy in public, which relates to our question because it helps conceptualize how people may support or oppose inferences because of their data source or destination [222]. People can have privacy concerns about public records in some contexts (inappropriate flows)

but not others (appropriate flows). In their study, participants express the greatest concern for privacy violations when data brokers were the sender of the information. Participants believe data brokers have increased capability to extract sensitive information. The idea that sensitive information can change hands without an observable flow increases the privacy risk.

4.2 The NIST Privacy Framework

Cybersecurity compliance alone being insufficient to protect privacy, NIST published its Privacy Framework [43] to help organizations with their privacy risk management programs. No statute or regulation requires framework adoption. The Privacy Framework's authors based its structure on the Cybersecurity Framework [223] but do not intend for it to be a privacy assessment or compliance tool [43]. Fifty-seven of 100 privacy subcategories are the same or similar to a cybersecurity subcategory [224], and both frameworks use the same list of mapped controls [46], [165]. NIST acknowledges that checklist-like requirements lend themselves readily to assessments that focus on compliance (not the framework's purpose) rather than on "achieving a positive outcome for privacy" [42, p. 14]. Although the NIST Privacy Framework [43] has achieved substantive goals for some adopters [44], it insufficiently identifies and mitigates inference risks.

4.2.1 Framework Structure and Implementation

The NIST Privacy Framework emphasizes communicating throughout an organization. It encourages dialogue among executives, managers, and practitioners to organize, assess, plan, and execute a privacy program in any organization, customized to that organization's needs [43]. The Privacy Framework [43] describes organizations' risk posture using profiles, a core, and implementation tiers. The idea is for organizations to iterate through the framework's core subcategories in their profile, and select applicable mapped controls for each. They can self-assess their competency with the implementation tiers.

The framework encourages organizations to generate *profiles* to model their privacy activities and risk tolerance: one representing their current posture and a target profile modeling their desired privacy risk-management goals. Organizations can use these profiles to self-evaluate their requirements against the posture. Each profile maps to the core's functions,

categories, and subcategories. The *core* embodies the highest levels of an organization's privacy activities as functions: identify, govern, control, communicate, and protect. Each function is divided into overarching, programmatic-level goals called categories and then technical- and management-level goals as subcategories. Each subcategory maps to various controls in the NIST SP 800-53 [46] as suggestions for mitigating the risk to that subcategory. We diagram this structure in Figure 4.3. Organizations can assess their own capability to manage privacy risk using the Framework's *implementation tiers*, that is, a progression of privacy program proficiency levels. The four tiers are partial, risk informed, repeatable, and adaptive [43].

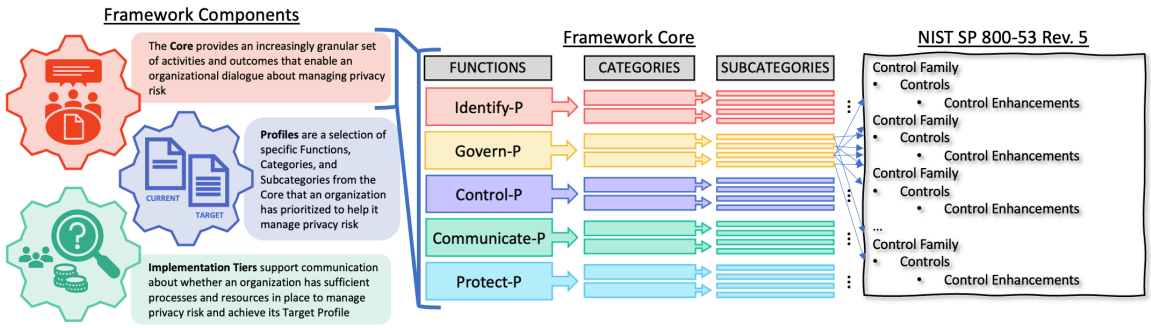


Figure 4.3. The NIST Privacy Framework's Structure, including Mappings from Subcategories to Controls. Adapted from [43, pp. 2 and 6].

NIST published revision 5 of SP 800-53 in 2020 to include both cybersecurity and privacy controls [46]. It groups hundreds of security and privacy controls into 20 families such as "Planning (PL)." Each family has base controls, each of which may have control enhancements identified by parenthetical suffixes. For example, AU-16(3) means "Audit and Accountability (AU) Control #16 (Enhancement #3)" and has the name "Cross-Organizational Audit Logging | Disassociability" [46]. Control names are succinct and do not by themselves sufficiently describe their controls. NIST maps most framework subcategories to specific controls and control enhancements [165]. Because mapped controls are suggestions, each adopting organization must choose controls for the organization's privacy goals. Selecting a control enhancement implies selecting its base control as well. Each control description can mention related controls and citations to help practitioners choose the best controls. Organizations must balance between applying every mapped control and not selecting an ap-

plicable control [61]. The subjectivity necessary to implement the NIST Privacy Framework well makes using it as a compliance or certification framework challenging.

4.2.2 Inferences in the Framework

Guarding against inference incidents is complex, and the framework should help. Protecting privacy is a sociotechnical endeavor. The framework applies to exposure of private data affecting individuals, "singly or in groups (including at a societal level)" [43, p. 4]. Organizations will be most effective in implementing any control (and the whole framework) as a whole-of-organization effort with cognizance of interactions between the system and component levels. The NIST Privacy Framework's only mention of "inference" is in the Control (CT) Function, Disassociated Processing (DP) Category, Privacy Subcategory #3 (P3) (CT.DP-P3).

> **CT.DP-P3:** Data are processed to limit the formulation of inferences about individuals' behavior or activities (e.g., data processing is decentralized, distributed architectures). [43, p. 24]

This definition applies to operational inferences and is about attributes rather than reidentification. Because many initially think of reidentification as the primary inference risk, the NIST Privacy Framework is of limited help.

Subcategory CT.DP-P3 maps directly to nine controls and control enhancements, listed in Table 4.1 [165]. Although most controls focus more on technical than organizational

aspects, three mapped controls (PL-8, PM-7, and SA-17) concern high-level actions that management can use to bolster the organization's privacy risk management.

Table 4.1: Controls from NIST SP 800-53 [46] Mapped [165] to Privacy Framework [43] Subcategory CT.DP-P3 and Our Summaries of Their Descriptions. Adapted from [46].

Control Family	**Control Number**	**Control Name**
		Summary
Access Control	AC-23	Data Mining Protection
		Limit database queries and responses. Apply differential privacy or homomorphic encryption and accountability notifications.
Audit and Accountability	AU-16(3)	Cross-Organizational Audit Logging | Disassociability
		Implement privacy-enhancing cryptography to disassociate individuals' identities from audit information.
Identification and Authentication	IA-8(6)	Identification and Authentication (Non-Organizational Users) | Disassociability
		Decrease identity attribute visibility to transmitting parties.
Planning	PL-8	Security and Privacy Architectures
		Develop them to describe requirements, dependency assumptions, and enterprise architecture integration.
Program Management	PM-7	Enterprise Architecture
		Develop and maintain the architecture to relate security and privacy risk with the organization's mission and business processes.

Continued on next page

Table 4.1: Controls from NIST SP 800-53 Mapped to Privacy Framework Subcategory CT.DP-P3 and Our Summaries of Their Descriptions (Continued)

Control Family	Control Number	Control Name / Summary
System and Services Acquisition	SA-8(33)	**Control Name:** Security and Privacy Engineering Principles | Minimization **Summary:** Minimize collection, processing, and retention of private data.
System and Services Acquisition	SA-17	**Control Name:** Developer Security and Privacy Architecture and Design **Summary:** Have external system developers produce a design specification and security and privacy architectures.
System and Communications Protection	SC-2(2)	**Control Name:** Separation of System and User Functionality | Disassociability **Summary:** Store applications separately from users' state information and interactions.
System and Information Integrity	SI-19	**Control Name:** Deidentification **Summary:** Remove elements of private data from data sets and evaluate reidentification viability.

Note: Each control number is serialized, for example, AU-16(3) means "Audit and Accountability (AU) Protection #16 (Enhancement #3)" and its name has the format "Base Control Name | Control Enhancement Name."

The framework addresses reidentification in subcategory CT.DP-P2 but focuses on data processing to "limit the identification of individuals" [43]. It does not address inferences directly. Of its eight mapped controls, three also map from CT.DP-P3 (AC-23, SA-8(33), and SI-19) [165]. The other five controls relate to reidentification not necessarily involving inferences.

4.3 Identifying and Mitigating Inference Risks with the Framework

To understand the value of the framework for mitigating inferences we assess how well it could have helped mitigate the incidents described in Section 4.1.1. These four inference incidents occurred before NIST published its Privacy Framework [43]; therefore, the involved organizations could not have used the framework. This forms a baseline from which we evaluate how well the NIST Privacy Framework could have helped in mitigating inference risk.

Applying Matland's ambiguity-conflict model of policy implementation, we assess the circumstances in the organizations before the inference incidents occurring to capture the nature of their privacy policy implementations, recognizing the incident as an outcome of implementation (see Figure 1.3) [9]. We recognize that the organizations involved were subject to legal privacy requirements. Although the specified requirements of these laws and regulations vary in their ambiguity, they all aim to protect people's privacy. Across the organizations, we find varied microimplementations enabled by the ambiguity in how to protect privacy. Although implementation conflict appears to be relatively low, points of conflict arise in balancing between data utility and privacy. For example, Strava's heat map, which is a product feature and depends on having users' location data, is at odds with protecting users' privacy (see Sections 4.3.3 and 4.3.8). Based on the overall conditions of high ambiguity and low conflict, the ambiguity-conflict model indicates that policy will be implemented when sufficient autonomy in microimplementation exists (see Figure 2.1) and that "contextual conditions dominate the [implementation] process" [9, p. 165]. We see in the incident descriptions in Section 4.1.1 how implementation occurs in varied ways, according to the context of each situation and organization, thus validating the model's relevance.

This section assesses of the effectiveness of the NIST Privacy Framework subcategory CT.DP-P3 and SP 800-53 controls to which it maps. Table 4.1 lists these controls and their enhancements. For each control and its enhancements, we assess whether it applies to each incident and would help identify and mitigate risk. We also estimate the control's likelihood of implementation. We summarize of our results in Table 4.2. Each cell's symbol in the middle four columns represents how much the control was implemented: **N/A** - not

applicable, **U** - unknown and unlikely, **L** - unknown and likely, **N** - not implemented, **P** - partially implemented, or **F** - fully implemented. The last column indicates the lesson observed: **CPAI** - combining anonymous data with public information, **PII/QI** - the insufficiency of removing solely PII categories or quasi-identifiers (QIs) to protect individuals from reidentification, or **TPTI** - difficulty translating policy into technical implementation.

Table 4.2. A Summary of Our Analysis of the Effectiveness of NIST SP 800-53 Controls [46] Mapped [165] to Privacy Framework [43] Subcategory CT.DP-P3 Against the Incidents Presented in Section 4.1.1

Control Number and Name	**EdX**	**NYC Taxi**	**Strava**	**Pizza Index**	**Lesson(s) Observed**
AC-23 Data Mining Protection	N	N	N/A	N/A	N/A
AU-16(3) Cross-Organizational Audit Logging | Disassociability	N	N	N/A	N/A	N/A
IA-8(6) Identification and Authentication (Non-Organizational Users) | Disassociability	P	P	P	P	PII/QI
PL-8 Security and Privacy Architectures	L	U	F	F	TPTI
PM-7 Enterprise Architecture	L	L	P	P	TPTI
SA-8(33) Security and Privacy Engineering Principles | Minimization	N	P	F	N/A	CPAI
SA-17 Developer Security and Privacy Architecture and Design	N/A	N/A	N/A	N/A	N/A
SC-2(2) Separation of System and User Functionality | Disassociability	P	N	P	P	PII/QI
SI-19 Deidentification	P	P	P	P	CPAI, TPTI, & PII/QI

Modeling this policy-implementation gap has multiple complications. First, we have limited insight into these organizations' policies to protect against inferences. Second, these organizations could not have used the NIST Privacy Framework. Third, we describe various uncertainties in this section, including in Table 4.2's U and L symbols. As such, our analysis represents an approximation based on the evidence and documentation available to us. Using a model to quantify these controls' implementations (like Equation 1.2 offers) can feign undue confidence in the apparent precision of our assessment [225]. Because of our limited ability to assess conformance of implementations to policy requirements (see Figure 1.3), using a model that spans goals to outcomes may be more appropriate.

4.3.1 AC-23: Data Mining Protection

This control says to limit database queries and responses and apply differential privacy or homomorphic encryption and accountability notifications. As indicated in Table 4.2, no organization implemented this control. Applying differential privacy well would have safeguarded both the EdX and taxi reidentification inference incidents; Cohen and Tockar both claimed as much [160], [183]. Limiting database queries or enabling accountability notifications for atypical database queries or accesses (as in [226] for multilevel relational databases or [151] for statistical databases) would not have helped because the data sets were made public. Applying homomorphic encryption, a technique suggested by this control, would not have helped in either incident because the plaintext data were public, and publishing encrypted data would have had no value for these organizations' intent to release the data.

Because Strava executed an aggregate query on its database, it abated data mining opportunities for reidentification in the heat map, but operational inferences were still possible. No PETs suggested by this control would have helped prevent the operational inferences. Similarly, for the Pizza Index, management deliberately revealed the inference publicly during a newspaper interview [189]. This control is therefore not applicable for operational inferences.

AC-23 has no control enhancements.

4.3.2 AU-16(3): Cross-Organizational Audit Logging | Disassociability

AU-16 focuses on the protection of users' identities when coordinating auditing with external organizations, like linking tables through blinded keys to disassociate logs from their data. An *audit* is an "independent review and examination of records and activities to assess the adequacy of system controls, to ensure compliance with established policies and operational procedures" [46, p. 395]. From a privacy perspective, to *disassociate* the data in logs means to remove fields and content that could link log entries with individuals. Although the principles of this control's "disassociability" enhancement could apply, despite no organization implementing it, auditing does not specifically concern the incidents in Section 4.1.1.

This control enhancement says to implement privacy-enhancing cryptography to disassociate individuals' identities from audit information. Privacy-enhancing cryptography (introduced in Section 4.1.4) may have helped mitigate reidentification inference incidents. Specifically, private information retrieval techniques could help in both the EdX and taxi incidents, if, for example, it provided a query interface for researchers instead of making the whole "anonymized" data set public. This control would not have helped mitigate the operational inferences because the incidents did not involve disclosure of individuals' identities; therefore, it does not apply to the operational inferences.

4.3.3 IA-8(6): Identification and Authentication (Non-Organizational Users) | Disassociability

IA-8 focuses on organizations' external, non-organization users' interactions with the organization's information systems. Because each incident described in Section 4.1.1 involves publicly released data, non-organizational users do not authenticate to access the data. This base control does not apply; though, the principles of IA-8(6) could still apply.

This control enhancement says to decrease identity attribute visibility to transmitting parties. Interestingly, each organization in the inference incidents partially implemented this control, in that they sought to "make [individuals' identifiers] less visible to transmitting parties" [46, p. 145], but their deidentification attempts were ineffective at mitigating the inferences. This is more evidence demonstrating the ineffectiveness of protecting privacy only by PII categories, which we describe in Section 4.4.3. In the EdX data, the quasi-identifiers

(QIs) remaining after deidentification contributed to analysts' reidentification inference capability. For the taxi data set described in Section 4.1.1, combining public information with released deidentified data enabled passenger reidentification. Strava's and the (perhaps inadvertent) Pizza Index's disassociation of individuals' identities through aggregation (like trying for k-anonymity with sufficient k) retained enough attribute information to enable these operational inferences. Neither Strava nor the Domino's franchise owner categorized geolocation as an identity attribute to be disassociated.

4.3.4 PL-8: Security and Privacy Architectures

NIST SP 800-53 describes security and privacy architectures as system-level manifests having three components: requirement descriptions for how each system protects the confidentiality, integrity, and availability of organizational information and PII, a description of how they support the enterprise architecture (explained in control PM-7 in Section 4.3.5), and assumptions about the system and its dependencies to accomplish these things [46, p. 198]. For our inference incidents, an architecture proved insufficient to mitigate them, indicating organizational difficulty with translating intent into policies and policies into technical implementations.

For reidentification inferences, the data owner's security and privacy architecture's implementation protects the data. While a private organization's system-level security and privacy architecture may not be publicly available, EdX had a privacy policy dated February 6, 2012 that was still in effect when the 2013 academic year began [227], the year of data comprising the EdX data set. The policy comprehensively covered the entire student experience, including that private data may be revealed publicly, "to the extent permitted by FERPA" [227]. As such, EdX likely had a system-level security and privacy architecture from which they derived this privacy policy, indicating apparent compliance, yet reidentification inferences were still possible. Similarly, New York City's law governed the Taxi Commission's handling of PII. Specifically, Chapter 5 of Title 10 of the administrative code defines PII and lists requirements for publicly disclosing that a security breach occurred [228]; however, in 2013, no statute required city agencies to have privacy policies or architectures and no public evidence that the commission had them exists. Therefore, the commission likely had some level of awareness of its need to protect individuals' privacy but likely lacked a security and privacy architecture. In this analysis, a gap in translation clearly exists. Without

additional internal details from these organizations, we speculate that the breakdown occurred in translating intent into policies, which would indicate the polices were insufficient, or translating policies into technical implementation.

For operational inferences, PL-8 differs from the other controls described thus far because the victim organization's architecture protects the data, not so much the data owner's architecture. In the cases of both Strava and the Pizza Index, the U.S. DoD (generalized here from the "Coalition Forces" that would have been in Afghanistan in 2018) was the affected party, not Strava, the data owner, or any individual. This is like a side-channel attack in that an unexpected information flow enabled the disclosure of sensitive information [182, p. 870]. The DoD has a system-level security architecture [229] and privacy program [230], but neither has a role in preventing the inference of sensitive information, which the OPSEC program [20] governs. Apparently, the DoD OPSEC program was insufficient in translating these types of operational inference risks for DoD employees to consider.

Although subcategory CT.DP-P3 maps to the PL-8 base control, framework adopters can implement any of PL-8's two control enhancements:

(1) Defense in Depth

This enhancement focuses on applying levels of defense while developing and administering security and privacy architectures [231]. Because the EdX and NYC Taxi data sets were made public, that is, outside an organization's control, layering additional defenses would not have prevented these inferences. Defense-in-depth could have helped affect an OPSEC program, however, in the Strava and Pizza Index incidents by suggesting additional safeguards to limit the signals coming from these DoD sites.

(2) Supplier Diversity

This enhancement addresses the potential issues of a monoculture [232]. Although diversifying the suppliers of IT products would not have applied in these incidents, one can apply this control enhancement's principle by considering the Pizza Index incident: If take-out orders were balanced across food genres or planners met virtually from geographically dispersed locations, the Domino's franchises' owner might not have been able to make the same index.

4.3.5 PM-7: Enterprise Architecture

This control says to develop and maintain an enterprise architecture to relate security and privacy risk with the organization's mission and business processes. In contrast with PL-8, PM-7 is a higher-level control that includes security and privacy (among a host of other organizational concerns) when integrating systems that might each have their own security and privacy architecture. In the incidents described in Section 4.1.1, an architecture proved insufficient to mitigate inference risk, indicating organizational difficulty with translating policies into technical implementations.

Like a system-level security and privacy architecture, an organization's enterprise architecture may not be public knowledge. EdX did not indicate on its Web site that it had an enterprise architecture, but EdX probably had a general business model that would meet the spirit of this control. NYC's Web site included a 2016 job posting indicating that the city has an enterprise architecture [233], though being a big and disparate organization, this may be a mis-categorization. Even so, without access to these architectures, we cannot determine how well they account for privacy risks. In contrast, the DoD (the generalized victim of the operational inference incidents) has only a framework for developing architectures [234], [235]. Nevertheless, it mentions risk management and guiding security and information assurance requirements but not privacy [234] so we gauge this as a partial implementation. Without additional details from these organizations, we speculate whether the breakdown occurred in translating intent into policies or translating policies into technical implementation.

"Offloading" is PM-7's sole control enhancement. It recommends that organizations move all non-essential supporting functions to other entities (including organizational components and third-party providers) to separate the functions from critical systems and data. This is like the principle of *least privilege*, that "every program and every user of the system should operate using the least set of privileges necessary to complete the job," which limits the actions that could compromise the confidentiality, integrity, and availability of a system and its data [236, p. 1282]. This control enhancement relies on the implied assumption that the component or contracted organizations performing these functions would be experts and thus have better quality and more efficient security and privacy safeguards specific to those functions, thus decreasing the likelihood of privacy violation. While how much EdX offloads non-essential supporting functions is unclear, governments typically delegate or

contract many functions so it is likely that both NYC and DoD offloaded functions to some degree. For all the above, nothing indicates that offloading contributed to the incidents.

4.3.6 SA-8(33): Security and Privacy Engineering Principles | Minimization

This is the only control with an engineering perspective on security and privacy. SA-8 and its enhancements serve as an avenue through which organizations can include security and privacy *engineering* principles. It also applies a lens by which the framework reminds its adopters to operationalize these principles at all stages of a system's life cycle. SA-8(33) focuses on minimizing the collection, processing, and retention of private data, a concept we discuss further in Section 4.4.3.

EdX collected students' level of education, gender, and year of birth, each as optional fields and inferred a country of residence from the user's public Internet Protocol (IP) address [160]. EdX's privacy policy allows at least nine purposes for these data, of which nearly all relate to data analysis [227]. Being optional, none of these PII fields matter to enrolling in and completing a course, the core function of the EdX platform. As such, EdX did not apply this security control. As explained in Section 4.1.1, these extra PII fields enabled reidentification inferences by allowing the combination of quasi-identifiers (QIs) [160].

The NYC TLC partially implemented SA-8(33), its data containing unnecessary PII fields for drivers but not passengers. These fields enabled analysts to learn the annual income of reidentified taxi drivers [181]; however, minimizing passenger PII proved ineffectual at preventing reidentification because of combining public information (listed in Section 4.1.1) with precise time and location data in the data set. Interestingly, unlike normal taxi service, mobile-app-based ride-share services (e.g., Lyft and Uber) require passenger identification for payment and "other purposes," which raises the question of passenger privacy in these contexts.

For Strava, minimization by aggregation did not prevent operational inferences [187]. Even without PII present in or contributing to the heat map, operational inferences were still possible because of public information related to current events [188]. On the other hand, nothing indicates that Domino's stored any of its customers' PII beyond transaction com-

pletion; such PII would not have applied to minimize to mitigate the operational inferences anyway. But for the other three incidents described in Section 4.1.1, combining public information with the deidentified data sets enabled these inferences.

4.3.7 SA-17: Developer Security and Privacy Architecture and Design

This control guides adopting organizations in setting requirements for external system developers. Since we have no indication that any of the incidents in Section 4.1.1 employed or contracted external developers, this control does not apply for these cases.

4.3.8 SC-2(2): Separation of System and User Functionality | Disassociability

SC-2 focuses on separating user-level and system- or privileged-level functions. Given the public nature of the data sets and inferences in the incidents, this base control would not have applied. The principles of SC-2(2) could still apply, though. It says to store applications separately from users' state information and interactions. For the one organization that appeared to have implemented this control enhancement (i.e., EdX), remaining quasi-identifiers (QIs) were problematic for preventing inferences.

For reidentification inferences, EdX's privacy policy [227] clearly communicated that it tracks users' interactions with the Web site; however, course enrollments and the number of forum posts per course (both QIs) were the only interaction data included in the data set [160]. In other words, we suspect that EdX did not normally separate its users' interaction data (to not hinder internal data analysis) but did so partially to release a deidentified data set. The NYC taxi data owners did not implement this control. To arrive at this conclusion, we considered that drivers' system interaction state information to be the association of the specific driver with each route completed.

For operational inferences, Strava's and the Pizza Index data aggregation removed nearly all users' association with their state information from the released data; but the retained user data enabled revealing users' locations. In this light, some view location data as identifying data [167], [237]. For example, even when similar heat maps establish endpoint privacy zones to hide users' starting and ending points, an attacker can still discover these precise endpoints [238]. Thus, Strava partially applied SC-2(2). If the DoD OPSEC Program alerted

Pentagon employees to the risk, they could have omitted information from their pizza orders (i.e., delivery address) by taking-out or dining-in, especially if at a further location than the closest Domino's to the Pentagon. Following this practice, Domino's franchise owner would have had to incur a greater burden to associate the orders with the Pentagon.

4.3.9 SI-19: Deidentification

This control says to remove elements of private data from data sets and evaluate for re-identification viability. EdX, NYC, and Strava each partially implemented SI-19 but, as previously explained, deidentified data insufficiently to prevent inferences. On the other hand, the Domino's franchises' owner deidentified his customers but not their association with the USG, which, in isolation, is not necessarily a sensitive correlation. Arguably, though, exposing this operational inference was the whole purpose of his interview with the Washington Post [189].

Although subcategory CT.DP-P3 maps to the SI-19 base control, framework adopters can implement any of SI-19's eight control enhancements:

(1) Collection

Deidentify the data a priori by limiting the collection only to the necessary fields. We analyze this concept of minimization under control SA-8(33) in Section 4.3.6.

(2) Archiving

To protect data stored long-term, this control enhancement urges deidentifying data before archiving them such that private data (whose intended utility was temporal) are not archived. Data archiving did not play a role in the selected incidents because the goal was data release.

(3) Release

Deidentifying data before releasing it outside the organization is the fundamental goal; however, the selected incidents demonstrated that it is a challenging endeavor. See also our discussion of disassociability and control IA-8(6) in Section 4.3.3.

(4) Removal, Masking, Encryption, Hashing, or Replacement of Direct Identifiers

In the EdX data set, fields that could serve as direct identifiers (PII categories) were either removed or replaced before its public release [160] but enough information remained to comprise QIs. The NYC TLC hashed direct identifiers, but did not use a key or salt to the hash as the control recommends, which enabled taxi driver reidentification [183]. Strava removed direct identifiers through aggregation. We analyze the Pizza Index deidentification from the Domino's perspective with this base control and, from the Pentagon perspective, with control SC-2(2) in Section 4.3.8.

(5) Statistical Disclosure Control

This control enhancement applies to situations in which multiple versions of a data set enable analysts to infer attributes because of changes from one version to the next. For example, an analyst could capture images of the Strava heat map at various times to create a longitudinal data set and make inferences about paths appearing or disappearing. For the other incidents, no comparison over time is possible because only one version of each data set was released, so this control does not apply.

(6) Differential Privacy

This is one of the most promising PETs to mitigate against reidentification inferences. We briefly list the foremost challenges to implementing differential privacy in Section 4.1.4 and analyze the potential success for differential privacy with control AC-23 in Section 4.3.1. Including the concept of inference generation from lingering QIs (as existed in the EdX deidentification process [160]) into controls like SI-19(6) "Differential Privacy" would prompt organizations (when selecting and discussing controls) to mitigate this inference vulnerability.

(7) Validated Algorithms and Software

Many privacy preserving technologies are available that help guard against inferences in specific use cases. For example: First, Hasan and Fritz devise a method that protects students from having their genders inferred from their online coursework interactions [239]. Second, Zhang et al. develop a method to anonymize data generated by human-wearable devices, inhibiting analysts' ability to infer age, gender, height, and weight while preserving the data

well enough to categorize activities, for example, walking and running [240]. Of the selected incidents, none used a validated deidentification algorithm like either of these examples.

(8) Motivated Intruder

Analogous to a penetration test for an information system, which is covered by control CA-8 Penetration Testing, this control enhancement involves trying to reidentify individuals in deidentified data. Tests should define the resources available to the motivated intruder, including level of insider knowledge and system access. Nothing indicates that any party in the selected incidents applied this control.

4.4 Recommendations

Based on our analysis of NIST's Privacy Framework helping organizations mitigate privacy inferences, summarized in Table 4.2, we recommend that privacy programs include the following inference-related aspects. Because they complement the framework, we also propose how to include these inference-related aspects in the framework [43]. To better mitigate an attacker's ability to combine anonymized data with other information, we seek to increase the coverage of inferences in mapped controls in Sections 4.4.1–4.4.2. First, we examine existing inference-related controls that are not mapped to subcategory CT.DP-P3. Then, we recommend improving specific controls with inference-relevant verbiage and references and then mapping those controls to CT.DP-P3. In Section 4.4.3, we propose updating controls about PII to account for the obsolescence of solely safeguarding PII categories to protect privacy. Finally, we discuss challenges associated with translating legal requirements and policy into implementable technical solutions in Section 4.4.4. Although we recommend only small changes to the framework's controls and their mappings to subcategory CT.DP-P3, making inference risk much more visible will improve organizations' risk awareness and ability to mitigate problems.

Because no inference incident we analyzed could have occurred had data had not been released, one may suggest that the simplest recommendation is to avoid releasing any data. The motivations for releasing data are informative when considering this recommendation. EdX released the data to gain research value. The Taxi Commission released the data to comply with a Freedom-of-Information-Law request. Strava released the data as part of its

business model. We do not know the circumstances that caused the Domino's franchises' owner to release the data. Because conflict exists between increasing the utility of data by releasing them and protecting the privacy of people by safeguarding them [154], [179], [241], frameworks like the NIST Privacy Framework should help organizations achieve a balance between the organization's goals and the privacy desires of the people whose data the organizations have.

4.4.1 Inference-Related Controls not Mapped

The NIST SP 800-53 has some inference-related controls hat are not mapped to CT.DP-P3 [46], [165], two controls containing words derived from "infer" and four other inference-related controls. For the two controls, we analyzed SI-19(5) in Section 4.3.9 and the other control, PL-4(1) "Rules of Behavior | Social Media and External Site/Application Usage Restrictions," would have helped mitigate the Strava operational inference. It relates to social media users' interactions with the organization's sensitive information. For example, if deployed coalition forces had considered their location to be sensitive organizational information (being in a foreign combat zone), then they could not, by policy, share their tracked workouts [187]. An OPSEC risk analysis could determine whether this information was safe to release. PL-4(1) is not mapped to any NIST Privacy Framework subcategory [165] but should be mapped to CT.DP-P3.

Although many controls in the NIST SP 800-53 [46] can help privacy, these four controls could help with privacy inferences:

AC-4(9) Information Flow Enforcement | Human Reviews Sometimes people can predict and mitigate the potential for inferences, as was the U.S. Census Bureau's standard practice before adopting differential privacy [150]. In the NYC taxi incident, people with the requisite expertise possibly could have foreseen the reidentification of taxi drivers and reidentification of passengers' identities.

AC-21 Information Sharing AC-21(1) "Automated Decision Support" could help with AC-4(9) if an ML-based PET existed to automate or help with human reviews by predicting a probability of inference. AC-21(2), "Information Search and Retrieval," relates to the query interface suggested with control AU-16(3) in Section 4.3.2 and differential privacy, discussed in Section 4.1.4 and Section 4.3.1.

SC-38 Operations Security Nearly all organizations have sensitive information to conceal from others. Although both the Strava and Pizza Index incidents relate to military organizations, private organizations can also protect their sensitive information with operational security (OPSEC) concepts. An OPSEC program is concerned with safeguarding specific non-sensitive information that could help an adversary discover sensitive information. National Security Decision Directive (NSDD) 298 summarizes the OPSEC process [170, p. 1]:

1. Identification of critical information
2. Analysis of threats
3. Analysis of vulnerabilities
4. Assessment of risks
5. Application of appropriate countermeasures

The directive also describes the importance of employee training to increase awareness of the organization's critical information and how to protect it against potential threats to better assure the organization's mission. Organizations may also direct their members to implement individual OPSEC controls, like using virtual private networks (VPNs) to obfuscate communications from eavesdroppers or not sharing fitness tracking data when at work.

SR-7 Supply Chain Operations Security SR-7 applies the principles of SC-38 to organizations' supply chains. It advises:

> Determining indicators that potential adversaries might obtain that could be interpreted or pieced together to derive information in sufficient time to cause harm to organizations . . . and considering how aggregated information may expose users or specific uses of the supply chain. [46, p. 370]

With Pentagon employees using Domino's as part of their unofficial supply chain, SR-7 might have helped mitigate this inference.

We recommend mapping NIST Privacy Framework subcategory CT.DP-P3 to each of the above four controls.

4.4.2 Other Controls for Addressing Inferences

Besides inference-related controls that were not mapped to CT.DP-P3, we found other controls that have a less direct inference relevance. Specifically, we recommend improving controls that concern organizational literacy of inferences, privacy violation disclosures, and organizations' privacy policies. We also suggest creating one new control to help organizations identify and mitigate the potential for privacy inferences.

Though many proprietary methods for threat models exist, the model of Linking, Identifying, Non-repudiation, Detecting, Data Disclosure, Unawareness, Non-compliance (LINDDUN) specifically applies to privacy and accounts for inferences [242]. Its name represents the threat types that can affect privacy. In 2024, MITRE offered Pattern and Action Nomenclature Of Privacy Threats In Context (PANOPTIC™) as a privacy threat modeling process [243]. Its structure and function is like other MITRE-published frameworks, including Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK®) [244]. To achieve this in the framework, we recommend bolstering control PM-16 "Threat Awareness Program" to include inference-related threats. From the discussion section of PM-16's description, we this italicized insertion:

> Because of the constantly changing and increasing sophistication of adversaries, especially the advanced persistent threat (APT), it may be more likely that adversaries can successfully breach or compromise organizational systems *or infer private organizational information*. [46, p. 211]

Then, PM-16 will support mapping CT.DP-P3 to a new control enhancement to base control AT-2 "Literacy Training and Awareness," which currently has six control enhancements [46]. This proposed seventh control enhancement, shown in Figure 4.4, would address inferences, including how they can occur and how to mitigate their potential harm. Increasing awareness among an organization's members requires knowledge, leadership, and persuasion to influence people's behavior [30], [97], [245], [246].

(7) LITERACY TRAINING AND AWARENESS | INFERENCES

Provide literacy training on recognizing adversarial opportunities for conducting reidentification and operational inferences of sensitive organizational information.

Discussion: Numerous inference-based incidents pepper recent history and have harmed organizations and individuals. Minimizing PII is insufficient for mitigating privacy inferences as motivated adversaries use all capabilities within their means to reidentify individuals in deidentified data and infer organizations' sensitive operational information.

Related Controls: AC-4, AC-22, AC-23, AC-24, PM-16, SC-38, SI-19

Figure 4.4. A New Control Enhancement for Control AT-2 in the NIST SP 800-53 [46]

One might argue that increased prevalence of PETs, like differential privacy, would decrease the need for organizational awareness of privacy inferences. Such an argument neglects the lessons of automation (described in Section 3.1), that automating a process can cause people to forget about the process and lose their ability to provide proper oversight of the now-automated process [86], [123]. The U.S. Census Bureau found the opposite circumstance, that implementing differential privacy required greater understanding among organizational personnel and census consumers of the privacy-enhancing methods [150], [217].

Increasing organizations' literacy on inferences enables a more effective execution of other controls. For example, having a greater understanding of inferences would enable people to include that understanding when implementing controls like AC-1 "Policy and Procedures," PL-8, PM-7, PM-28 "Risk Framing," and RA-8 "Privacy Impact Assessments," to which multiple framework subcategories map [165]. Inference training is especially key for the personnel involved in executing controls AC-22 "Publicly Accessible Content" and AC-24 "Access Control Decisions" because of their role in publicizing data from which threat actors could make harmful inferences.

These efforts can also lead to including inference concepts into organizational privacy program documentation. For example, an organization can increase its privacy transparency

by publishing a privacy policy that clearly addresses the sources used in making inferences and communicates, as suggested by Wachter and Mittelstadt [202]:

- The organization's intent to infer sensitive data from non-sensitive data.
- The organization's intended purpose for the inference and how the source data are relevant, which should align ethically with their use.
- The assumed and proven accuracy and reliability of the organization's inference methods.

Control PM-20(1) "Dissemination of Privacy Program Information | Privacy Policies on Websites, Applications, and Digital Services" addresses this, stipulating that organizations' privacy policies be relatively easy to understand and enable informed consent [46]. "Notice and consent" is a common legal standard [33] but opponents claim that transparency about an organization's use of personally sensitive data is more important. Transparency rather than forcing consent to privacy policies would lead to "inaccessible [privacy-exposing application programming interfaces (APIs)], coding to prevent scraping, automatic deletion of data, and blocking of cookies" [33].

Besides increasing awareness of inferences, organizations' sensitive information disclosure procedures for incident response also need modification to account for inference incidents. Therefore, we also recommend expanding the following disclosure-related controls with our *italicized* additions and ~~strikethrough~~ deletion to include inference concepts.

AU-13 Monitoring for Information Disclosure "Examples of organizational information include personally identifiable information retained by the organization or proprietary information generated by *or inferred about* the organization."

PM-21 Accounting of Disclosures "Develop and maintain an accurate accounting of disclosures of personally identifiable information *and incidents of reidentified individuals in purportedly anonymized organizational data*."

RA-8 Privacy Impact Assessments "A privacy impact assessment is an analysis of how personally identifiable information *and sensitive organizational information* ~~*is*~~ *are* handled to ensure that handling conforms to applicable privacy requirements, determine the privacy risks associated with an information system or activity, and evaluate ways to mitigate privacy risks." [46]

Other NIST Privacy Framework subcategories map to these controls [165], which help the organization-wide privacy preserving effort advocated by the framework itself [43].

4.4.3 PII Protection Requirements

As our understanding of the importance and limitations of protecting PII category data grows, the NIST Privacy Framework should evolve accordingly. Even though protecting PII categories is insufficient for protecting privacy, we recommend updating all controls in the NIST SP 800-53 [46] that depend on identifying PII categories when determining what data to safeguard. These updates are in the same spirit as Holder's recommendations to update the GDPR and CCPA language for inferences [198, pp. 1352–1353]. These would also improve framework subcategory CT.DP-P2, described in Section 4.2.2.

The NIST SP 800-53 lists 83 controls in that mention PII, listed in Figure 4.5 [46]. We found 4 (5%) that have wording sufficient to include inference concepts. For example, PT-7(2) says to prohibit the processing of any information describing how specific people exercise their First Amendment rights. Also, 17 (20%) controls mention PII simply as an example of sensitive information to organizations and associated people. For example, AC-4(1) describes information flows containing *sensitive information, including PII* as one example. We do not recommend any change to these controls to account for inferences because they indicate that PII categories are not the sole privacy concern. Similarly, 22 (27%) controls are adequate as they are because they describe safeguarding PII, which is still important, without depending directly on identifying and protecting its categorized data as the sole method of privacy protection. For example, SI-12(1) says to *limit* the use of *PII* elements *to help reduce* the level of *privacy risk*. Finally, 40 (48%) controls need modification to incorporate inferences. For example, SI-19(3) says to remove PII from a data set before its release. In these controls, appending "and other sensitive information that could lead to reidentification of individuals" to instances of "personally identifiable information" would be minimally sufficient.

Incorporates Inferences	A Type of Sensitive Information	Safeguarding PII	SI-18(3)	Needs to Incorporate Inferences	PT-4
AC-23	AC-3(11)	AC-2	SI-19(1)	AC-16	PT-4(1)
PL-4(1)	AC-4(1)	AC-3(14)		AT-2	PT-4(2)
PT-7(2)	AC-4(17)	AU-3(3)		AT-3	PT-4(3)
SI-19(7)	AC-21	AU-9		AT-3(3)	PT-5
	AU-2	AU-12(4)		AT-3(5)	PT-5(1)
	AU-3	AU-14		CM-13	RA-3
	AU-13	PE-8(3)		IR-2(3)	RA-8
	MP-6	PL-8(2)		IR-4	SA-8(33)
	PM-23	PM-22		IR-8	SA-15(12)
	PM-26	PT-2		IR-8(1)	SC-7(24)
	RA-5(5)	PT-2(1)		PL-2	SC-16
	SA-3(2)	PT-2(2)		PL-8	SC-42(4)
	SA-11(5)	PT-3		PM-5(1)	SI-12(2)
	SI-7	PT-3(1)		PM-9	SI-18(2)
	SI-11	PT-3(2)		PM-11	SI-18(4)
	SI-12(3)	PT-7		PM-20(1)	SI-18(5)
	SI-20	PT-7(1)		PM-21	SI-19
		SI-12(1)		PM-25	SI-19(2)
		SI-18		PT-1	SI-19(3)
		SI-18(1)			SI-19(6)

Figure 4.5. Controls Mentioning the Phrase "Personally Identifiable Information" in the NIST SP 800-53 [46]. No control contains the acronym "PII."

One can conclude that the privacy regulations that endorse k-anonymity for deidentification are now obsolete yet they remain in force as the last standards to have been approved by policymakers.

4.4.4 Translating Legal Requirements into Technical Implementations

As a society, we generally value privacy; however, many concepts of privacy exist, including what it is and entails, and diverse opinions about sharing, trust, obfuscation, decisions to

control one's privacy, and protection from those that might intrude into or exploit one's private matters [33], [247]. Legislatures often define principles and standards in laws rather than defining clear rules; whereas the former is often not computable in logic standards, computers apply rules very well [17], [248]. This distinction matters because organizations must translate legal requirements into policies for the laws to have their intended effect. In turn, organizations must translate those policies into technical implementations. Strategic ambiguity and delegation of detail in laws provides flexibility in implementation [9]. Elliott et al. provide a comprehensive method and framework of considerations [154] based, in part, on the United Kingdom Anonymization Code of Practice [249].

Many controls in the NIST SP 800-53 [46] mention the term "law" or "legal" in contexts unrelated to translating law to the organizational policy or technical levels. Most uses of "law" refer to complying with "applicable laws" or referencing "law enforcement." Most instances of "legal" recommend seeking legal counsel for matters addressed by a control [46]. In consideration of these usages, we do not recommend any modification to the NIST Privacy Framework or controls to improve organizations' ability to translate law into policy and technical implementations.

Instead, legislatures and organizations can help overcome translation challenges, which could reduce litigation in the judiciary. Legislatures can work with scientists to establish a formal, mathematically provable definition of privacy or just implementable requirements [161], [248], [250]. People with the International Association of Privacy Professionals (IAPP)'s Certified Information Privacy Professional (CIPP) certification (which exists for "putting privacy law and policy to work" [251]) may benefit both legislatures and organizations. Unless AI can replace much of the tedious labor in the legal profession, improvisational narrative and discourse to address cases not previously examined need human involvement because they have no precedent in the training data [252]. As such, systems at all levels (legislative, policy, technical, and judicial) that implement legal logic standards must adapt to new laws or else risk obsolescence in a "technological–legal lock-in" that stifles legal evolution [253] and, to a lesser extent, industry.

A concept called *checklist compliance* can arise easily within organizations trying to meet legal or policy requirements. It involves believing that completing a checklist verbatim will achieve the checklist's objective. In analyzing controls AU-16(3), IA-8(6), SA-8(33), and

SC-2(2) (in Sections 4.3.2, 4.3.3, 4.3.6, and 4.3.8, respectively), we describe how applying the principle of a control is most effective compared to its verbatim implementation, which is like checklist compliance. For example, Strava ostensibly implemented controls IA-8(6) and SC-2(2) by considering that identity attributes are the same as PII category data (a checklist of specific data fields) instead of approaching identity from a contemporary perspective that recognizes location as an identity attribute [167], [238]. Privacy impact assessments (PIAs) provide another checklist compliance opportunity as organizations try to comply with applicable laws and government regulations; they help protect organizations by their mere existence as a paper trail but, by themselves, do not provide security or privacy [33].

Checklists have their place and, when developed and implemented well, can be highly useful for guiding professionals through most any complex process. Gawande comprehensively describes how creating an effective checklist is a challenging endeavor, specifically in achieving the right level of detail for the people that will be executing the checklist [93]. NIST recognizes that frameworks lend themselves readily to checklist-style assessments that focus on compliance rather than on "achieving a positive outcome for privacy" [42, p. 14]. Although one can construct a checklist by listing every phrase in a standard and putting a checkbox next to it for indicating it has been addressed, such an approach ignores the need for professional judgment in assessing the quality of each checklist item. For example, control PM-20(1) (mentioned in Section 4.4.2) describes the characteristics of an organization's public-facing privacy policy, like "written in plain language and organized in a way that is easy to understand and navigate," which is subjective and requires professional judgment to assess [46, p. 214]. The EdX data were treated according to FERPA deidentification requirements but reidentification was still possible [160]. While we can argue that FERPA deidentification requirements are outdated, deidentification and reidentification techniques will continue to evolve, quickly making any new deidentification checklist outdated. This is why checklists for complex tasks like deidentification, flying airplanes, and surgery should contain principles for professionals to execute, guided by their professional judgment [89], [93].

Every organization must apply the framework uniquely, according to its goals and privacy requirements [76]. An appropriate checklist for the NIST Privacy Framework might include general goals, like communicate within the organization, identify risk, mitigate risk, and

repeat continuously. Such a checklist could serve as a risk communication tool to help organizations achieve their goal of reducing privacy incidents.

4.5 Conclusion

As organizations face growing complexity in protecting privacy and risk in using sensitive data, resources to help organizations identify the nature and scope of their privacy risk must evolve. The practical need for organizations to comply with privacy laws and meet people's privacy protection expectations led us to ask how we could improve the NIST Privacy Framework [43], an important guide to implementing organizational privacy programs. Frameworks like the NIST Privacy Framework should help organizations achieve the balance between data utility and people's privacy. We focus on mitigating inference-based privacy violations, taxonomically defining inferences as reidentification or operational inferences.

How can standards or frameworks guide or hinder organizations' policy implementation? To determine how to improve organizations' defenses against privacy inferences using the NIST Privacy Framework, we apply the framework to past incidents of reidentification and operational inferences. This analysis revealed shortcomings in the framework's capacity to help organizations identify inference risk. Moreover, when organizations use the framework as a checklist (against NIST's instructions), their success depends on the professional judgment of privacy and cybersecurity practitioners who are poised to recognize and overcome its shortfalls. In other words, every adopting organization requires a unique application of the framework.

Matland's ambiguity-conflict model of policy implementation explains how the local autonomy needed for privacy risk management enables the varied implementations of privacy policy across organizations [9]. Our recommendations include increasing organizations' awareness of inferences by expanding the mapping of inference-related controls and improving selected other controls to account for inferences, updating controls that depend on removing specific PII categories or quasi-identifiers (QIs) as sufficing for protecting privacy, and improving organizations' ability to translate legal requirements into policy and policy into technical implementations. Further analyses of NIST Privacy Framework effectiveness would contribute to this field of research, especially if done by framework-implementing organizations on privacy incidents or near-incidents involving inferences.

CHAPTER 5:
Implementing an Enterprise Cybersecurity Policy across the U.S. Federal Executive Branch

Beyond adopting modern information technology and using standards and frameworks to improve their software systems, organizations' own structural complexity can affect the policy-implementation gap. Within large organizations, like the U.S. Government, having diverse and disparate components managing cybersecurity can be difficult. Our research question asks, "What can cause the implementations of a mandatory cybersecurity policy to vary across U.S. Federal Executive Branch organizations?" We use Matland's ambiguity-conflict model (introduced in Section 2.2) to understand the nature and origin of the policy-implementation gap in this context [9], which enables setting cybersecurity policies that reduce the gap [52].

We examine .gov domains owned by U.S. Federal Executive Branch agencies, which is our population for this study. Our focus is on exploring factors affecting their conformance with the organizational cybersecurity policy to implement the specified controls. Exploring differences in the policy-implementation gap across the branch provides insight into enterprise cybersecurity policy implementation and compliance. We conduct an observational study in which the U.S. Government (USG) issues a top-down cybersecurity policy directing its components, approximately 100 agencies, to adopt specific technical controls. This study shows the importance of understanding the influence of *organizational factors* in cybersecurity-policy implementation across the U.S. Federal Executive Branch [14].

The organizational structures that affect how policies and their implementations translate between goals and outcomes (see Figure 1.3) influence the variations we observed: Each agency must interpret the same policy requirements and implement the same controls under the same security model but using possibly different assessment tools and compliance regimes. Even so, agencies have similar access to cybersecurity tools, techniques, and resources. U.S. law provides a mechanism for the Cybersecurity and Infrastructure Security Agency (CISA), a Department of Homeland Security (DHS) component, to specify requirements for the cybersecurity programs of some USG agencies through Binding

Operational Directives (BODs) and Emergency Directives (EDs) [254]. Under this authority, a publicly-accessible policy, BOD 18-01: "Enhance Email and Web Security" [19], specifies configuration requirements to U.S. Federal Executive Branch agencies. The policy promises worthwhile cybersecurity gains ("ensure the integrity and confidentiality of internet-delivered data, minimize spam, and better protect users" [19, p.1]) and appears simple to implement (see Section 5.1.1), yet non-conformance remains substantially high (reported in Section 5.2.1). Understanding the reasons for their non-conformance reveals challenges with implementing an enterprise cybersecurity policy in the USG.

We measured BOD 18-01 conformance by querying Internet-visible endpoints we discover using domains listed in the public .gov zone file published by CISA. These initial measurements serve as both motivation for and subjects of this study. Our population consists of the 1,190 Federal Executive domains in the .gov zone from 126 agencies, of which 582 (49%) domains from 99 (79%) agencies remained deficient six years after the original implementation deadline. Considering just the 97 agencies over which CISA claims authority (see Section 5.1), 468 (44%) domains from 71 (73%) agencies remain deficient. We used our gap measurement model from Section 1.1 to devise a measure for comparing configurations' *distance to conformance* with the specified policy. When we found a nonconforming service, we sent a request for information (RFI) to the domain's listed contact, asking for reasons for their non-conformance. Agencies' reasons for nonconforming include misunderstanding the controls in the policy, resource priorities, forgetfulness, questions of authority and responsibility, and operational reasons. More recent public-policy efforts such as Executive Order 14028, "Improving the Nation's Cybersecurity" [255], and the National Cybersecurity Strategy [256] increase the need to couple cybersecurity policy to practice better.

In Section 5.1, we describe the observational study of BOD 18-01 conformance in the U.S. Federal Executive Branch. We describe our data collection work in Section 5.2 and analyze the policy-implementation gap in Section 5.3. We discuss the implications of our findings in Section 5.4 and conclude in Section 5.5.

5.1 BOD 18-01, An Observational Study

The U.S. DHS issued BOD 18-01 in October 2017, directing Federal Executive Branch agencies to implement these measures [19]:

1. For email security, within 90 days:
   (a) Configure valid Sender Policy Framework (SPF) records on second-level domains.
   (b) Configure valid Domain-based Message Authentication, Reporting, and Conformance (DMARC) records on second-level domains with a minimum policy of `p=none` and at least one address specified for a reporting uniform resource identifier (URI) for aggregate data (RUA) or reporting URI for failure data (RUF).
   (c) Enable start Transport Layer Security (STARTTLS).
2. For Web server security, within 120 days:
   (a) Enable Hypertext Transfer Protocol (HTTP) Strict Transport Security (HSTS).
   (b) Disable Secure Sockets Layer (SSL).
   (c) Disable Triple Data Encryption Standard (3DES).
   (d) Disable Rivest Cipher 4 (RC4).
3. For the DMARC records:
   (a) Within 15 days of the National Cybersecurity and Communications Integrations Center (NCCIC) being established, specify NCCIC as an RUA.
   (b) Within one year of BOD 18-01, set a reject policy, `p=reject`.

In our study, BOD 18-01 [19], the CISA operational directive enabling statute [254], and the Federal Information Security Modernization Act (FISMA) [54] function as the public policies affecting each agency's organizational policy (referring to Figure 1.2). U.S. Federal Executive Branch agencies are required by law [54] and regulation [257] to set up internal cybersecurity governance programs conforming to requirements that lean heavily on guiding documents published by the National Institute of Standards and Technology (NIST), including the Federal Information Processing Standards (FIPS) and Special Publications (SPs) in the 500, 800, and 1800 series. Each agency's FISMA-driven cybersecurity management plan serves as the organizational policy. Their standardized structure, including cybersecurity assessment practices, enables comparing policy implementation across covered agencies,

despite their independent governance and goals. The BOD 18-01 requirements that we measure serve as the implemented policy.

This creates an observational study: The same policy applies across many organizations with approximately comparable cybersecurity governance and assessment approaches, each of which is required by law to conform to the policy where conformance is directly and non-invasively measurable from the public Internet. The agency is the independent variable varying widely in mission, structure, and employees and the implementations offer a set of dependent variables. As such, we can observe the effects of variability in organizational structure on cybersecurity risk perception and management as proxied through implementation.

CISA maintains the list of second-level .gov domains and their registration metadata in which we find 1,190 domain names registered as "Federal – Executive" by 126 agencies [258]. Other government entities in the U.S. may register .gov domains, but BOD 18-01 legally applies only to Federal Executive Branch agencies, according to the BOD authority in 44 U.S.C., Subchapter 35 [254]. Exceptions include statutorily defined National Security Systems (NSSs) and certain other defined Department of Defense (DoD) and Intelligence Community (IC) systems. CISA lists 102 agencies for which the BOD authority is legally required [259]. The 126 agencies choosing "Federal – Executive" when registering their domains that are not among these 102 are mostly independent commissions, federally chartered corporations, and DoD/IC agencies. Of these 102 agencies, 4 are components of another of the remaining 98 agencies. Also, one agency registered its domain under the General Services Administration (GSA), despite not being a GSA component. Thus, we have a sample of 97 unique agencies subject to the authority of BOD 18-01.

Next, we describe BOD 18-01's requirements. Then we describe our method in Section 5.1.2 and our distance-to-conformance model in Section 5.1.3.

5.1.1 BOD 18-01 Requirements

While BOD 18-01 specifies configuration parameters for DMARC, it simply requires enabling SPF, STARTTLS, HSTS, and disabling SSL, 3DES, and RC4 [19]. The policy specifies that the SPF and DMARC records be "valid" but does not define the standard for determining a record's validity. The description of our methods (Section 5.1.2) discusses the

concept of record validity. We summarize BOD 18-01's SPF and DMARC requirements' hierarchy in Figure 5.1.

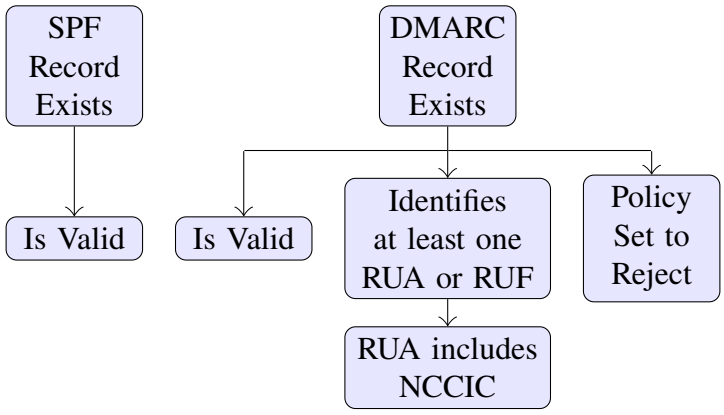


Figure 5.1. Conformance Status Tree for BOD 18-01's SPF and DMARC Requirements. The arrows show the order of implementation, so a child node can have a conforming status only if its parent conforms.

SPF

SPF records are Domain Name System (DNS) text (`TXT`) records that list all servers authorized to send email for that domain [260]. Upon receiving an email, mail transfer agents should verify (using the "Return-Path," not the "From" field) that the email came from a server authorized by the sender's domain's SPF record. SPF employs qualifiers (+, -, ~, and ?) and mechanisms (e.g., `all`, `include`, `redirect`, `a`, and `mx`) to shorten SPF record length. When an SPF record lists a name instead of an Internet Protocol (IP) address or has an `include` mechanism (which includes the senders allowed by the included domain's SPF record), the mail transfer agent must again query DNS to resolve it. The '-' qualifier preceding the `all` mechanism yields a "fail" status for all email servers not listed before it in the record. The `all` mechanism has an implicit '?' (neutral) qualifier, which can permit all servers to send email on behalf of the domain [260]. An oversized allowed-IP address space in an SPF record could allow attackers to spoof a domain [261].

Durumeric et al. found (in 2015) that 40% of the Alexa top million domains have SPF records [262], improving to 49% in 2018 [263], and 56.5% in 2023 [264]. Foster et al.

analyzed in 2015 SPF implementations from both the originating domain and email receiver perspectives [265], finding that 85% of email providers have SPF records. We build on prior measurement work by investigating factors hindering U.S. Federal Executive Branch agencies' adoption of SPF and DMARC, even when a legitimate authority directs their use.

DMARC

DMARC records are DNS text records on the "_dmarc" subdomain [266]. If an email fails an SPF or Domain Keys Identified Mail (DKIM) (mentioned in but not directed by BOD 18-01) check, receiving email servers should query the domain's DMARC record for disposition (i.e., whether to *reject*, *quarantine*, or not disrupt delivery (*none*)) and reporting instructions. A best practice that the BOD recognizes is starting with "none" or "quarantine" for monitoring the DMARC reports. Once administrators have ensured correct functionality, they can set a "reject" policy, thus avoiding the email delivery denial-of-service (DoS) that an unconfirmed record with a reject policy could cause.

The reporting URI for aggregate data (RUA) and reporting URI for failure data (RUF) fields respectively specify where to send DMARC aggregate and failure reports and take the form of a comma-delineated list of `mailto:` email address URIs. If any email address's destination domain differs from the queried domain, the destination domain must have a "Report Receiver" DNS text record indicating that it accepts DMARC reports from the queried domain [266]. *Aggregate reports* include summarized metadata on the emails a server receives from the subject domain over a specified duration, including whether they passed the SPF and DKIM checks. BOD 18-01 requires agencies to list the NCCIC reporting email address in the RUA field of each domain's DMARC record [19]. Such centralized reporting enables CISA's cross-domain analysis across the .gov enterprise. Receiving email servers generate and send *failure reports* upon a DMARC failure. BOD 18-01 does not specify any failure reporting options (FO) configuration, which prescribes whether SPF, DKIM, either, or both must fail for the receiving email server to generate a failure report for sending to the RUF.

SPF and DMARC Working Together

SPF and DMARC together help prevent unauthorized impersonation of their registered domains, as they did for the New Jersey Office of the Attorney General in 2023 [267]. This

helps agencies protect their reputation from harms that could come from impersonation, including loss of public trust and lost revenue in recovering from the incident. When they are not in place, an attacker can spoof a .gov domain as the originator of phishing emails, as happened in 2021 [4]. Often, cybersecurity proponents recommend combining SPF, DKIM, and DMARC with other email security measures [268], [269]. Because BOD 18-01 does not direct DKIM implementation, DKIM is beyond the scope of this chapter.

SPF and DMARC work as intended only if receiving servers use them. In 2015, Foster et al. found that 10 (45%) of 22 popular email providers enforce a reject policy to comply with originating domains' SPF and DMARC records [265]. By 2023, DMARC enforcement improved to 25 (53%) of 47 popular email providers [269]. Some researchers find that SPF and DMARC configurations are difficult to form [270]. Hu et al. find that a critical point occurs just short of 100% adoption, at which the benefit of adopting SPF and DMARC overcomes the cost of implementation [271]. Internet routing security faces a similar challenge of achieving critical points for adopting routing security advancements. Industry experiences moderate success in overcoming this challenge through the Mutually Agreed Norms for Routing Security (MANRS) initiative [272], but no similar initiative for email security exists.

STARTTLS

In the Simple Mail Transfer Protocol (SMTP), STARTTLS is a keyword that can initiate a Transport Layer Security (TLS) handshake, but does not itself promise any security properties [273]. STARTTLS studies generally show gradual progress toward better security and identify measures to overcome STARTTLS weaknesses [262], [274].

HSTS

HTTP Strict Transport Security (HSTS) enforces secure-mode connectivity with Web sites by using an HTTP header that must be sent over a secure connection. Once recognized as an HSTS-enabled server, a client will expect to connect by TLS and Hypertext Transfer Protocol Secure (HTTPS), and should close the connection when HTTPS negotiation fails. HSTS helps organizations that frequently handle sensitive information, effectively forcing clients to connect to servers by TLS. This mitigates the potential for attacks, like man-in-the-middle (MitM) [275].

SSL

The Secure Sockets Layer (SSL) protocol is one of the earliest transport-security protocols for securing network traffic online. Early iterations faced problems, such as a lack of data sequencing and integrity checks, prompting a series of security updates, including SSL 2.0 and 3.0. Subsequent updates use the protocol name TLS and add new features while eliminating vulnerabilities. The Internet Engineering Task Force (IETF) deprecated SSL in 2015 in favor of TLS [276, Chapter 2].

3DES

3DES, also called TDES or Triple Data Encryption Algorithm (TDEA), is a variant on the Data Encryption Standard (DES) block cipher [277]. It improved upon previous versions by enciphering each block three times using independent secret keys. Ultimately, however, it was shown to be vulnerable to ciphertext collision, revealing exploitable information about corresponding plaintext blocks [278], prompting NIST in 2019 to disallow 3DES in TLS [279].

RC4

RC4 is a stream cipher, once adopted worldwide, including in Secure Shell (SSH) and TLS [280]. It produces a ciphertext by XOR operation of the plaintext and a pseudorandom sequence of bits. As a non-FIPS-approved algorithm, RC4 was temporarily allowed in government systems when a FIPS-approved encryption algorithm was unavailable [281]. Cryptanalysis has since shown exploitable vulnerabilities, leading NIST to withdrawal its limited status in 2014 [282] and the IETF to deprecate it in 2015 [283].

5.1.2 Measurement Methods

What has hindered agencies from fully conforming with BOD 18-01, a seemingly simple cybersecurity policy, in six years since the directive's issuance? We began by measuring BOD 18-01's requirements. We expanded our understanding of these measurements with qualitative survey data revealing explanations for conformance variation. Finally, we measured conformance again to capture how it changes with time before conducting our analysis.

We measured BOD 18-01 conformance by querying public servers based on CISA's list of registered .gov domains [258]. We used `trustymail` to query DNS for measuring SPF and

DMARC conformance [284]. `trustymail` checks only for an SPF or DMARC record's validity according to its specification in RFC 7208 [260], not its semantic correctness. Additionally, if a DMARC record needs a report-receiver record, `trustymail` verifies these external destinations and, if missing, marks the DMARC record as invalid. We also used `trustymail` to probe the servers listed in the mail exchange (MX) DNS record for measuring STARTTLS conformance. The probing consists of sending `elho` or `helo`, interpreting the received response or timing-out, and sending `quit`. We provide these data in Supplemental 4. To verify Web-server security, we used the portion of Glade's data collected on .gov domains [285].

Although we measure domains' conformance, the agencies are responsible for domains' conformance. The distribution of domains registered and owned by agency is wide, as shown in Figure 5.2: Most agencies have one or two registered domains but could own 126 domains (mode 1, median 2, population mean 9.6, and population standard deviation 20.8).

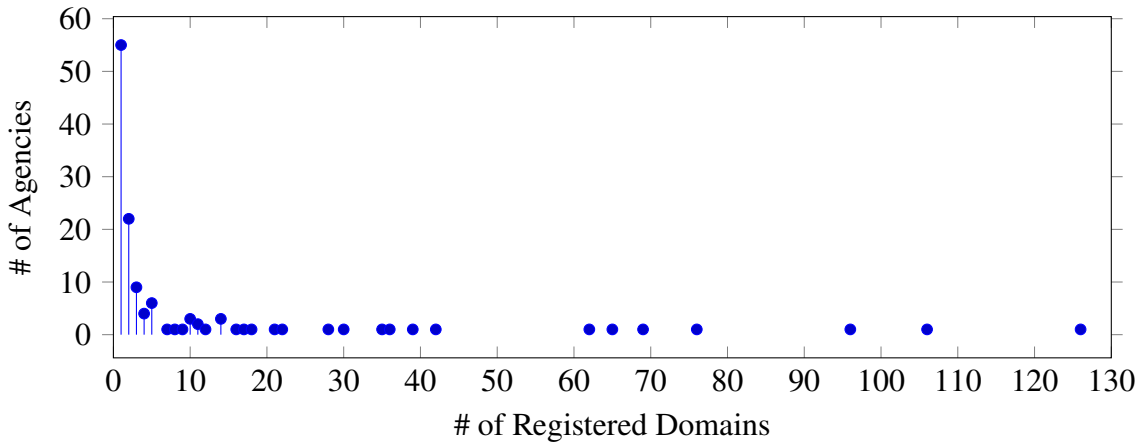


Figure 5.2. Domain Registrations per Agency

We survey the nonconforming agencies by sending them RFIs, according to our template in Appendix C. Three months after sending our RFIs, we re-queried DNS to measure the changes in SPF and DMARC conformance. Our three-month timeline matches the initial 90-day conformance deadline specified in BOD 18-01 [19]. We used our distance-to-

conformance model to measure domains' degree of conformance with the SPF and DMARC requirements measured in our first and second queries.

5.1.3 Distance to Conformance

Given the specific policy requirements and their corresponding controls listed in BOD 18-01, combined with our ability to measure their implementation for each domain, this study is ideal for using a model (like that offered in Equation 1.2) to quantify policy implementation. According to the procedure provided in Section 1.1, once we have calculated $cip(I, P)$ for each domain (Step 3), we can average them to calculate an aggregate $cip_A(I, P)$ for the population of Federal Executive Branch agencies' domains (Step 4). Because of the policy requirements' hierarchical nature, we make the following adaptations. We quantify the possible statuses in conformance status tree, T (shown in Figure 5.1), by assigning each node, n, a value of 1 if nonconforming and 0 if conforming. Thus, we measure *distance to conformance*, d, by adapting an inverse of our gap measurement model in Equation 1.2, that is, $1 - cip(I, P)$ in which 0 is conforming and higher values are less conforming. To have d represent the number of implementation steps to conform rather than a ratio or percentage to complete, we omit the denominator. One can subscript the d with the size of the tree, $|T|$, to clarify the implementation's progress; in this chapter, the conformance tree remains static at a size of 7 (as shown in Figure 5.3), so we omit this subscript notation for simplicity.

$$d_{|T|} = \sum_{n \in T} n \tag{5.1}$$

Figure 5.3. A Representation of the Possible Outcomes per Domain of the Distance-to-Conformance Model, a One-Dimensional Number Line Ranging [0–7]

Because the policy requirements are *hierarchically related* (as shown in Figure 5.1, none of the descendent nodes can have a conforming status until their parent node has a conforming

status), we summarize their values by assigning a monotonically incremented sum of a node's child nodes' values. For example, the absence of a DMARC record has a value of five because its node in T has four descendent nodes. Practitioners must take five, partially hierarchically dependent, implementation steps, as delineated in the policy, to cover the implementation steps from a "No DMARC Record" status to a fully conforming DMARC configuration.

Our distance-to-conformance model is like a weighted edit distance, computable as a Hamming distance [286], except that our model compares configurations only with a conforming configuration, $d = 0$. *Hamming distance* is a metric that represents the number of non-matching bits between two equal-length strings, that is, the count of 1s in an XOR operation's output. Using the same "No DMARC Record" example, we can represent the conformance status tree as a string of bits, `0011111`, which has a Hamming distance of 5 with the conforming status, `0000000`. In its role as a comparator, we can see that one configuration is farther from conforming than another; however, unlike Hamming distance, it does not afford a meaningful comparison directly between two nonconforming configurations because the conformance model is not a metric.

An important consideration is that, while a comparison of d values indicates relative conformance levels, it does not necessarily enable a comparison of the cybersecurity severity implications of non-conformance. For instance, an invalid SPF record ($d = 1$) is closer to conforming than not having an SPF record ($d = 2$). But from a security perspective, an invalid record is likely worse because this condition can inadvertently permit any server to send an email for the domain (including malicious servers). The initial building block of a means to measure policy conformance, however, remains our focus in this chapter.

We believe this measure is generalizable to measuring other hierarchical conformance requirements. We compared it with other measurement models and metrics that quantify an implementation's degree of conformance to a cybersecurity-related policy. Because many models are created for a specific purpose, they may not be reasonably adaptable to a different use case, like representing BOD 18-01 conformance. We found one other model with a similar goal of quantifying an implementation's degree of conformance to a policy, that is, a model for assessing NIST Cybersecurity Framework compliance [287]. That model assesses an organization's implementation's conformance with the framework's mapped controls

across three vectors against a five-dimensional capability maturity model and with three-dimensional weights to incorporate relative importance. Its use requires that organizations identify their levels of cybersecurity maturity for human resources optimization, information technology (IT) capabilities, and processes in addition to the relative importance (weight) of each framework subcategory. As such, it is more complicated than is necessary for representing BOD 18-01 conformance. Our distance-to-conformance model is simpler and is likely more flexible to adapt to other policies.

5.2 Data Collection

Consistent with our method (described in Section 5.1.2), we describe our initial query and server probing measurements in Section 5.2.1. In Section 5.2.2, we describe the statistical tests that we use in subsequent sections. To supplement our initial measurements, we describe our qualitative survey in Section 5.2.3 and a quantitative longitudinal conformance status update in Section 5.2.4.

5.2.1 Initial Measurements

By agency, conformance rates are higher for every BOD 18-01 requirement among the subset of agencies over which CISA claims authority than all Federal Executive agencies in our population. By domain, one exception inhibits the same observation: STARTTLS conformance is slightly lower on CISA-claimed agencies' email servers than the population of all .gov email servers. Because the conformance rates are relatively similar between all agencies' domains in our population and the subset of agencies under CISA's authority (Δ_{max} = 9.3%), we use the population for our analyses and use relevant samples as specified. In the following subsections, we explain more of these details measured by querying DNS and probing servers, respectively.

SPF and DMARC Measurements by DNS Queries

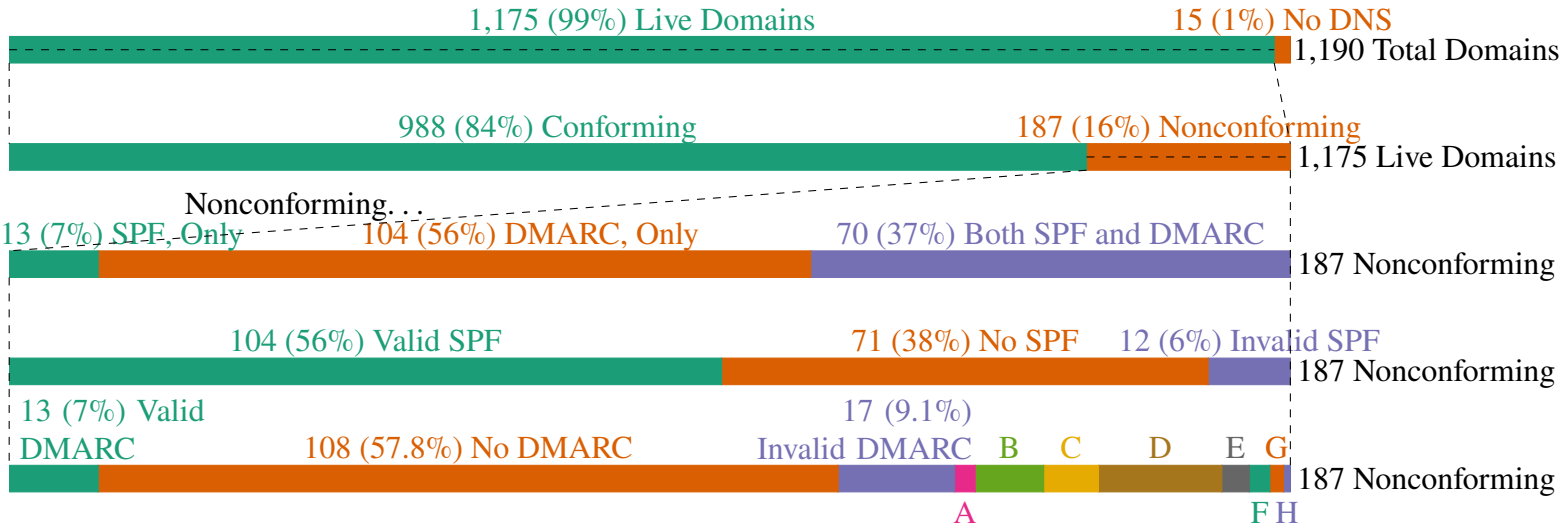
According to our measurements in July 2023, 988 (83%) conforming of 1,190 domains belong to 75% of agencies, and the nonconforming domains (17%) belong to 41% of the agencies. All domains owned by 74 (59%) of 126 agencies meet the SPF/DMARC requirements of BOD 18-01.

`trustymail` labels domains having public DNS records as *live*. We look first at whether each registered domain is live, then for BOD conformance, grouping nonconforming domains according to the requirements not met. If a domain is not live, we assign it a distance-to-conformance, $d = 0$. Because the .gov registry does not also offer DNS hosting, agencies must acquire DNS hosting services after registering a .gov domain, limiting the registry's ability to enforce policy directly.

Using the conformance tree in Figure 5.1 as a guide to BOD 18-01's hierarchical requirements, we find in our measurement data various configurations of nonconforming domains. We identify conformance status such that if a requirement (node) is not met, then the node's children's statuses are unable to be met. For example, if a domain lacks an SPF record, the SPF record's syntax is irrelevant, so we report only the lack of an SPF record and not its invalid syntax. Only about half of the nonconforming domains have a valid SPF record (104 of 187, 54%) but less than one-tenth have a valid DMARC record (13 of 187, 7%). Table 5.1 shows the number of domains that we measure in each nonconforming configuration and their respective distance-to-conformance values. Figure 5.4 offers an alternate perspective to explain these details. For example, one can readily sum Table 5.1 rows having the blue "No SPF Record" to see that it matches the "No SPF" value in Figure 5.4, $64+5+1+1 = 71$.

Table 5.1. How Domains Fail to Conform with SPF and DMARC Requirements and Their Distance-to-Conformance Values. *n* represents the count of domains and *d* represents each nonconforming configuration's distance to conformance.

Nonconforming Requirement(s)		***n***	***%***	***d***
No SPF Record and No DMARC Record		64	34.2	7
No DMARC Record	(Total 108)	41	21.9	5
Reject Policy Not Set and NCCIC not in RUA		18	9.6	2
Invalid DMARC Record Syntax	(Total 23)	16	8.6	1
Reject Policy Not Set	(Total 36)	9	4.8	1
Invalid SPF Record Syntax	(Total 12)	8	4.3	1
NCCIC not in RUA	(Total 29)	8	4.3	1
No SPF Record	(Total 71)	5	2.7	2
No RUA or RUF and Reject Policy Not Set		4	2.1	3
Invalid SPF Record Syntax and No DMARC Record		3	1.6	6
No RUA or RUF	(Total 7)	3	1.6	2
Invalid DMARC Record Syntax and Reject Policy Not Set		2	1.1	2
Invalid DMARC Record Syntax and NCCIC not in RUA		2	1.1	2
No SPF Record and Invalid DMARC Record Syntax		1	0.5	3
No SPF Record and Reject Policy Not Set		1	0.5	3
Invalid SPF Record Syntax, Invalid DMARC Record Syntax, and Reject Policy Not Set		1	0.5	3
Invalid DMARC Record Syntax, Reject Policy Not Set, and NCCIC not in RUA		1	0.5	3
Total		**187**	**100**	**-**



A: 3 (1.6%) No RUA or RUF
B: 10 (5.3%) Reject Policy Not Set
C: 8 (4.3%) NCCIC not in RUA
D: 18 (9.6%) Reject Policy Not Set, NCCIC not in RUA
E: 4 (2.1%) No RUA or RUF, Reject Policy Not Set
F: 3 (1.6%) Invalid DMARC Record Syntax, Reject Policy Not Set
G: 2 (1.1%) Invalid DMARC Record Syntax, NCCIC not in RUA
H: 1 (0.5%) Invalid DMARC Record Syntax, Reject Policy Not Set, NCCIC not in RUA

Figure 5.4. BOD 18-01 SPF and DMARC Conformance Categorization by Domain

Not having both SPF and DMARC records is the most frequent deficiency (115 of 187 nonconforming domains, 34 of 126 agencies), indicative of agencies not trying to conform rather than trying and falling short. Second, administrators have not been setting a "reject" policy in their domains' DMARC records (36 domains, 22 agencies), despite the one-year timeline that BOD 18-01 grants to ramp up the policy [19]. Of the domains' DMARC records lacking a "reject" policy, 23 have a "none" policy (15 agencies) and 13 have a "quarantine" policy (10 agencies). Third is not including the NCCIC address in the RUA (29 domains, 16 agencies). The fourth and fifth most common nonconforming requirements are invalid record syntax for DMARC (23 domains, 13 agencies) or SPF (12 domains, 5 agencies).

The cause of every invalid DMARC syntax error comes from a failure to verify the external domain of an RUA or RUF email address. When an RUA or RUF includes an email address in a different domain, that external domain lacking the requisite report-receiver text record invalidates the DMARC record. This query contains 1,051 domains with DMARC records' RUA or RUF URIs needing external verification of which 23 fail. This 98% external destination verification rate (1,028 of 1,051 Federal domains) is better than the 75% rate that Hureau et al. measure in a 2024 Internet measurement study of 5.4 million domains with an external RUA or RUF [270]. Although BOD 18-01 does not address external destination verification in its text, CISA subsequently described it in the implementation guidance on the BOD 18-01 Web page [19]. When the report-receiver record is missing, the DMARC specification says the domain receiving mail "MUST NOT generate a report" [266, Section 7.1 (emphasis in original)]. Two of these errors involve reporting domains with no MX records so no email service to receive reports is reachable by name.

Of the 12 SPF records with invalid syntax, we find 5 distinct errors [260]. Some records contain multiple errors:

1. Apparent `all` mechanism misconfiguration (5/12). `trustymail` does not check for this important aspect, negating SPF's purpose by allowing any server to send email for the domain.
2. An included domain not having an SPF record (4/12).

3. Resolution requiring too many DNS lookups (2/12). “Too many DNS lookups” is a syntax error because of the risk for DoS attacks against DNS. The lookup limit per SPF record is 10 [260].
4. Trivial recursion by a domain including itself (2/12).
5. Omitting Backus-Naur Form delimiters (spaces) (1/12).

Server Endpoint Security Measurements

Fifty-four (43%) of 126 agencies have an email server that does not support STARTTLS. Twenty-six of these agencies lack any conforming email server. Every agency has an email server. Of email hosting domains, 220 (41%) of 533 support STARTTLS.

Only 120 Federal Executive Branch agencies have public-facing Web servers on 1,021 registered domains. Most domains have Web servers at the second-level domain (883) while 13 domains have Web servers only on subdomains. The remaining 125 domains have Web servers on both the second-level domain and subdomains. Sometimes multiple endpoints serve a single domain’s Web services, as in content delivery networks (CDNs). Most Federal Executive Web servers are accessible by 1 or 2 endpoints (distinguished by IP address), but some have up to 27 endpoints (mode 1, median 2, sample mean 2.4, sample standard deviation 2.35). As such, we count a domain as conforming only when all its endpoints conform, including its subdomains. We have conformance measurements for 3,044 endpoints (2,170 unique) representing 1,259 Web servers in agencies’ second-level and subdomains. Because domains sometimes share endpoints, as in cloud-service providers and CDNs, a nonconforming endpoint can affect multiple domains’ conformance. Most endpoints (1,882 (87%) of 2,170) host a single domain but some endpoints (by IP address) host many domains, up to 31 (median and mode 1, sample mean 1.4, sample standard deviation 1.81).

Of the 120 agencies having a public-facing Web server, 71 (59%) do not conform with at least one policy (HSTS, 3DES, or RC4) in its domain or one of its subdomains. All conform fully with BOD 18-01’s requirement to not support SSL. Stratifying by agency, HSTS support looks bad, as 70 (58%) agencies own a domain that does not support HSTS, of which 23 agencies do not have any Web servers supporting HSTS; however, stratifying by domain reveals better HSTS support: 791 (77%) of 1,021 domains and their subdomains

conform. The number of domains per agency nonconforming with HSTS varies widely, 61 agencies having 5 or fewer domains without HSTS (range $[1 - 29]$, mode 1, median 2, sample mean 3.5, and sample standard deviation 5.12).

Ciphersuite conformance is comparatively high. Only 15 (13%) agencies have Web servers that support the deprecated 3DES ciphersuites on 20 domains (2% of 1,021). Of the 20 domains that support 3DES, 2 domains (owned by separate agencies) also support the deprecated RC4 ciphersuites. No other domains support RC4 ciphersuites.

5.2.2 Statistical Tests

We conduct five statistical tests on our measurements that yield p values, all Pearson χ^2 tests of independence. For four of the five χ^2 tests, we test for the independence between an independent variable (different for each test) and the dependent variable of whether the domain conforms to BOD 18-01's SPF and DMARC requirements. The fifth χ^2 test instead uses a dependent variable of STARTTLS conformance. To ensure that we are not unduly inflating the probability of rejecting a false null hypothesis, that is, controlling the family-wise error rate, we apply Holm's sequentially rejective multiple test procedure [288]. We report the adjusted p values in the text describing each of these tests. Separately, we calculate Spearman's Correlation Coefficient, r_s, on the population of agencies, which does not involve a p value. Supplemental 4 contains the calculations for our statistical tests and adjustments. We address our use of each statistical test and its assumptions in the text with the result of the test.

Two assumptions of the Pearson χ^2 test of independence warrant greater discussion before we use the tests. The first assumption is that the *observations* are independent, which is different than the independence of *variables* that the test evaluates [289, p. 152], [290, p. 659]. Our observations are the measurable configurations of our population of domains, which meet this test's assumption of independent observations. A formal definition of independence is $P(A \cap B) = P(A) \cdot P(B)$, that is, the probability that both A and B occur "is the product of their individual probabilities" because one does not depend on the other [290, p. 87]. In other words, our configuration measurements of domains A and B are independent because neither domain's configuration depends on the other's configuration,

even if the same person configured multiple domains or people across agencies collaborated to form their configurations.

The second assumption relates to the minimum expected values in a contingency table (cross-tabulation). All expected values must be greater than or equal to one and, as a *safe heuristic* to cover most cases, should be greater than or equal to five in at least 80% of cells to ensure an accurate type I error rate [291], [292]. A lack of consensus on the minimum expected values complicates determining acceptable uses of the test. Many have settled on a heuristic value of five in at least 80% of cells; but as a heuristic, it may be unnecessarily conservative. Sampling experiments revealed that the chi-square test remains robust (having error rates between 0.0187 and 0.0891 at $\alpha = 0.05$) with expected values of 1–5 under some circumstances [292]. The circumstances include having a sample size greater than or equal to 20 and that the data in the contingency table can be highly skewed. To reduce the type I error rate when the contingency table has low expected values, users can use a more conservative α. "Since the maximum inflation in error rate for $N \geq 20$ is less than double the nominal value, using an adjusted level of $\alpha/2$ should suffice in most cases for the provision of adequate protection at the original level desired" [292, p. 1,295]. Sometimes choosing a more conservative α will not affect whether one rejects the null hypothesis because the test's calculated p value is not close to α. Nevertheless, when one's "interest is in making a decision about the presence of association in the $R \times C$ table and not in estimating exact cumulative multinomial probabilities" [292, p. 1,296], one can bypass the heuristic, which requires having minimum expected values of one, $N \geq 20$, and $\alpha_{adjusted} \leq \alpha/2$. We address this aspect in the text as it arises in our tests.

We recognize the effect of collaboratively forming domains' configurations prior to us observing and measuring them as a baseline characteristic of the population of U.S. Federal Executive Branch .gov domains rather than as affecting the independence of our observations. The nature of this collaboration could combine directive, structural, or informal means. Formally, the guidance on BOD 18-01's website [19] is sufficient to claim that all agencies' SPF and DMARC configurations are dependent on the same source of guidance. Further, multiple domains owned by the same agency may be configured by the same person or team, according to the organization's structure. Moreover, unofficial collaboration may occur between personnel across agencies or simply by mimicking another agency's configurations. Other avenues that could lead to an informal collaboration in form-

ing configurations include administrators following guidance found in Web search results and interacting through the affordances of an agency's DNS provider's interface. While we can identify definite collaborations, we might be able to discern some other collaborations based on organizational structure but not others that remain unknown without extensive field studies.

5.2.3 Requests for Information

To discover factors hindering agencies' conformance with BOD 18-01's SPF and DMARC requirements, we sent the RFI in Appendix C by email to the 98 contacts associated with 187 nonconforming domains.

Like Hureau et al. [270], we implemented a procedure to find suitable contacts for these RFIs. Hureau looked for contacts in the order: RUA/RUF, WHOIS, and RFC 2142 aliases. We preempted this list with the .gov domain registration's "security contact" field. This field was blank for 94 (8%) of 1,190 domains. For 63 of these domains, we resolved 38 blank contacts with an addressee from another domain registered by the same agency. If there was more than one address, we chose the address that we suspected would most likely garner a response. We filled the remaining 25 addressees with those specified in RUAs.

In reviewing the measured conformance data, there appeared to be an unusually high number of nonconforming domains among those without registered security contacts. To test the hypothesis that a correlation exists between whether a domain's registration has a security contact (the independent variable) and the domain's SPF and DMARC conformance (the dependent variable), we construct a contingency table (cross-tabulation in Table 5.2) and use a Pearson χ^2 test of independence. The null hypothesis is that no correlation exists between these two variables [289, p. 147]. The assumptions of the test are as follows:

- The data are categorical [289, p. 139].
- Each domain contributes exactly one count to exactly one cell in the contingency table [289, p. 153], [291].
- The observations (measured domain configurations) are independent, as described in Section 5.2.2.
- We overcame the requirement for randomly selecting the sample by selecting the population [291].

- For the expected values (described in Section 5.2.2), this instance meets the heuristic because the smallest expected value for Table 5.2 is 14.

We find a statistically significant correlation between domains' conformance and whether they have registered security contacts, $\chi^2 = 203$ and $p = 5.00 \times 10^{-46}$, adjusted to $p = 2.50 \times 10^{-45}$ by Holm's procedure. In other words, we reject the null hypothesis and conclude, based on the expected and observed values, that agencies that are lax in completing the .gov domain registration form are very unlikely to conform with the SPF and DMARC requirements of BOD 18-01, a cross-contextual consistency observation [293].

Table 5.2. Cross-Tabulation of SPF/DMARC Conformance vs. Domain Security Contact Registration by Domain

	With Contact	No Contact	Totals
Nonconforming	124	63	(16%) 187
Conforming	972	31	(84%) 1,003
Totals	(92%) 1,096	(8%) 94	1,190

From 98 RFIs, we received 10 answers, 5 from agencies under CISA authority. Our 10% response rate is much better than recent large-scale email-security surveys [269], [294] but on par with a 2021 survey with manually formed recipient lists [295]. Another 2 responses indicate a willingness to answer but are working through an approval process; however, follow-up garnered no response. We also received the following 12 automated responses: 3 out-of-office notifications (we forwarded the RFI to identified forwarding addresses), 2 "thank you for contacting us" messages absent of any promise to follow up, 1 "automated ticket creation" notification that lacked a way to follow up, and 6 undeliverable email error messages. Our 6% undeliverable rate is much better than recent, large-scale notifications (including those reporting shortfalls in SPF and DMARC configurations) that depend on WHOIS [269], [270], [296]. We attribute 1 false positive to a `trustymail` or DNS anomaly; neither the recipient nor we could replicate the false positive. The remaining 73 RFIs did not yield a response.

The RFI responses offer qualitative insight into the factors hindering agencies' conformance with the BOD. Most notably, most respondents appear to be unaware of the purpose of SPF and DMARC. Some respondents have operational or resource limitations hindering their conformance; other respondents point to a lack of authority or confusion in responsibilities. Contributing to finding registered domains without live DNS records, we learn that some agencies have found .gov domain registration to be cumbersome, so when their use for a domain name ends, they may keep the name registered for later reuse.

The 10 agencies answering our RFI cover 26 of the 187 nonconforming domains. Three answers provide two reasons for non-conformance so the count of reasons in the list below appears to total 13. This list's numbering corresponds with the list at the end of Section 5.2.4. The frequency of each reason is as follows:

1. 6 answers claim that the subject domains do not need SPF and DMARC records because the domain has no email server, a confusion we address in Section 5.3.1. For example, "[Domain] is NOT an email enabled domain, so DMARC record is not required for this domain at this time" and "The [domain] domain is currently not used for email services." Despite having a live DNS record, another respondent submits its reason as, "[Domain] was sunsetted over 7 years ago."
2. 3 answers cite operational reasons for why they do not conform. One of these answers explains that many sources must send email for the domain, which explains the apparent misuse of SPF's implicit `?all` mechanism and lack of a reject policy. The other two answers claim that the domains are still in a quarantine period before escalating to a reject policy.
3. 2 answers relate to authority. One claims a lack of authority to implement the policies on a subset of the agency's registered domains, either directly or through a subordinate. The other cites DHS's/CISA's lack of authority to direct its agency (see Section 5.4.2).
4. 1 answer cites insufficient manpower, which translates to prioritization. The respondent elaborates, " 'Things that are not automated don't get done automatically'—let's call it 'manpower issues' or 'other reasons.' "
5. 1 answer blames the agency it contracts to manage its DNS records. In a follow-up email exchange between the author, the domain owner, and the contracted agency, the contracted agency explains that it would create SPF and DMARC records upon the domain's owner's request.

Two of these answers also report updating their configuration to conform, which we confirmed. For example, one respondent claims, "Thank you for your email and for bringing this issue to our attention. We've addressed the issue and the new DMARC records for both domains are now as follows. . . " We replied to all answers with a "thank you" note. For the 6 domains described first, we also included a short, non-confrontational description of SPF and DMARC (shown in Appendix D), explaining that they protect a domain against being spoofed as an email origin.

5.2.4 Conformance over Time

We re-queried agencies' DNS records in October 2023 to measure any conformance changes 90 days after transmitting our RFIs, to align with the DHS's initial 90-deadline [19]. Coincidentally, October 2023 marks the sixth anniversary of DHS publishing BOD 18-01, yet large gaps in conformance persist. In this section, we use our distance-to-conformance model to show domains' SPF and DMARC conformance changes from the first (d_1) to the second query (d_2).

Between queries, Federal Executive domain registration increased from 1,190 to 1,195 [258]. Agencies unregistered 8 conforming domains, no nonconforming domains, and 3 that were not live. Agencies registered 16 new domains of which only 6 conform with the SPF and DMARC requirements, 9 do not conform, and 1 is not live. Our combined domain list from both queries has 1,206 unique domains. If CISA enforced BOD 18-01 conformance as a condition of domain registration, one might expect that all newly registered domains would conform; however, as we explain in Section 5.2.1, CISA lacks its own DNS server on which to enforce its policies.

Of the 1,179 domains in both query data sets, nearly all (1,139) *maintain* their d ($d_1 = d_2$), 10 *decline* ($d_1 < d_2$), and 30 (including the anomaly described in Section 5.2.3) *improve* ($d_1 > d_2$) in their conformance. We show these transitions in Table 5.3 along with whether an RFI listed the domain, the agency received an RFI for another domain, or the agency did not receive an RFI. Consistent with Step 4 of the procedure provided in Section 1.1, averaging implementation measurements yields aggregates of $d_{1_A} = 0.6751$ and $d_{2_A} = 0.6277$.

Table 5.3. Domain Conformance Over Time versus RFI Status

	Domain in RFI	Another Domain in RFI	Agency not Contacted	Totals
Improved Conformance	29	0	1	(2%) 30
Maintained Conformance	150	525	464	(97%) 1,139
Declined Conformance	0	5	5	(1%) 10
Totals	(15%) 179	(45%) 530	(40%) 470	1,179

From an agency perspective, the changes are mixed, shown in Table 5.4. Further, agencies over which CISA claims authority own all 10 declining domains and 18 of 30 improving domains.

Table 5.4. Longitudinal Domain Conformance Changes, Grouped by Agency

Agencies			Domains			
Agencies Have Domains That. . .		Count	Improve	Maintain	Decline	Total
Improve, only	$(d_1 > d_2)$	5	6	-	-	6
Improve & Maintain	$(d_1 \geq d_2)$	12	23	364	-	387
Maintain, only	$(d_1 = d_2)$	103	-	522	-	522
Decline & Maintain	$(d_1 \leq d_2)$	4	-	179	6	185
Decline, only	$(d_1 < d_2)$	1	-	-	3	3
Improve, Maintain, & Decline		1	1	74	1	76
Totals		126	30	1,139	10	1,179

The 10 declining domains comprise 3.8% of the domains registered by six agencies, all under CISA's directive authority. Domains with declining conformance had $d_1 = 0$ and d_2 ranging $[1 - 7]$ with a median of 1, sample mean 1.9, and sample standard deviation 1.76, indicating that most domains non-conformed with one additional requirement but

a few outliers non-conformed with several requirements (shown in Figure 5.5). Notably, 5 of these 10 domains have registered contacts to which we sent an RFI for another of its registered domains. We counted only 1 of these 5 domains whose registered contact responded to the RFI: The agency's response is 1 of the 6 that we describe in Section 5.2.3 as claiming to have no need for SPF and DMARC because the domain has no email server. Despite replying with our SPF and DMARC explanation (Appendix D) as our "thank you" note for responding to our RFI, this domain declined from $d_1 = 0$ to $d_2 = 2$ because it had no SPF record in the re-query data set. Although a small ratio of domains for 5 of these 6 agencies (1.3%–7.1%), 3 of the 10 declining domains comprised all the domains registered by one agency. Interestingly, this agency only changed one of its domains' DNS records, dropping its DMARC reject policy and removing its "Report Receiver" record, which invalidated the other two domains' DMARC records without either of their DMARC records changing.

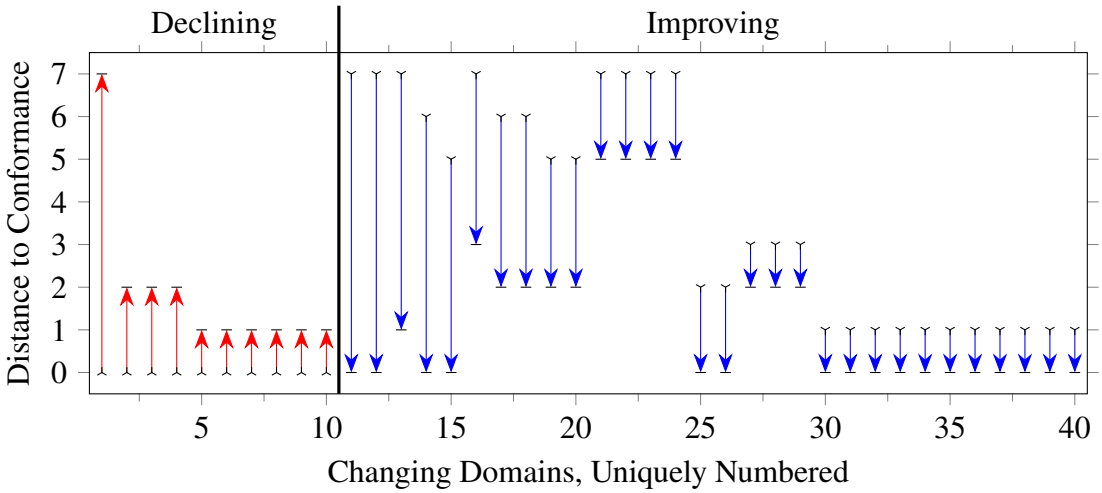


Figure 5.5. Changing Domains' Δds

All 30 domains whose conformance improved were subjects of our RFIs, except one whose agency did not receive an RFI from us. Like recent email security studies [264], [269], [270], we do not know whether any of the 29 contacted agencies' domains' conformance would have improved had we not transmitted any RFIs; our population of 98 agency contacts with a 10% response rate is not large enough to have implemented a meaningful

control group [294], [295]. The 30 improving domains comprised a minority (1%–7%) of 8 agencies' domains, a mid-range (17%–50%) portion of 5 agencies' domains, and totality of 5 agencies' domains. The 30 domains' d improved (i.e., decreased), on average more than those in the previous paragraph declined, Δd ranging $[1 - 7]$ with a median of 2, sample mean 2.5, and sample standard deviation 1.93 (shown in Figure 5.5). The improvements brought 6 of 18 affected agencies to full SPF and DMARC conformance.

Of the 10 RFI responses, those 26 domains either improved (6) or maintained (20) their conformance status. This list's numbering corresponds with the list in Section 5.2.3:

1. 1 of 6 contacts improved its domain's conformance from $d_1 = 7$ to $d_2 = 0$. One other contact removed all DNS entries associated with one of its domains so that it is no longer live. This contact's other domain and the other 4 of 6 contacts maintained their nonconforming status.
2. Both contacts that claimed their domains were in a quarantine period still had not set a reject policy.
3. For items 3–5, no status changes, except for one domain with its DNS entries removed, the same domain as that described in the first item above.

5.3 Analysis

We explore factors affecting agencies' relatively low conformance by integrating our query data and survey data and analyzing the situation through the ambiguity-conflict model of policy-implementation [9]. Then, we consider a risk-reward tradeoff for each of BOD 18-01's requirements and the nature of how the conformance of each requirement affects conformance with other requirements. In Sections 5.3.1–5.3.3, we form and test hypotheses based on our observations. Finally, we address effects of agencies outsourcing their domains' email services and form recommendations for improving conformance of outsourced domains in Section 5.3.4.

BOD 18-01 is a top-down policy from the U.S. Department of Homeland Security (DHS), an effort now led by CISA, directed at U.S. Federal Executive Branch agencies [19]. The policy's precision facilitates its low ambiguity. Even so, confusion about how the directed requirements work has caused some ambiguity. As our RFI responses in Section 5.2.3 revealed, some agencies have experienced conflict in implementing this policy, though, the

policy has relatively low conflict overall. This combination means that once an agency can resource the implementation, the policy's implementation is likely (see Figure 2.1). Our query and server probing data corroborate that.

The 10 RFI responses represent some edge cases that organizations faced. This list's numbering corresponds to that in Section 5.2.3:

1. Confusion about how SPF and DMARC work is a microimplementation ambiguity. Because domains without email service are more likely to conform to these requirements (explored in Section 5.3.1), it seems that the independence of each agency's cybersecurity team has facilitated autonomy in microimplementation. However, instead of being the catalyst for implementing policy, the autonomy enabled misunderstandings to exist in some agencies' cybersecurity teams. This is not a breakdown of the ambiguity-conflict model. Rather, "misunderstanding" is a reason for a low-ambiguity, low-conflict policy to not be implemented as intended [9, p. 161], implying that having people that understand the policy is a "resource."
2. The operational reasons for non-conformance reveal a microimplementation conflict, specifically, that one agency's mission depends on having many servers sending email for its domain. This means that conforming to the SPF and DMARC requirements would hurt its mission. We categorize this conflict as occurring at the microimplementation level because the CISA can handle it on a case-by-case exception basis. BOD 18-01 allows exceptions if agencies can mitigate the threat by other means: "Explain the challenges expected to cause the delay [in conforming] and how the agency plans to promptly overcome the challenges" [19, p. 3].
3. Respondents blaming a lack of authority for their non-conformance suggests organizational conflict. One organization experienced conflict at the microimplementation level in which it lacked authority to direct its own component. Another organization experienced conflict at the macroimplementation level, claiming that it is not subject to the BOD authority, an issue we address in Section 5.4.2. In either case, as conflict surrounding a policy increases, those with the greatest power to act on the policy effect the implementation (see Figure 2.1).
4. In implementing an overall low-ambiguity, low-conflict policy, agencies primarily need sufficient manpower resources, which identifies a factor hindering this agency's domains' conformance.

5. At first, this case may look like ambiguity because the agency is not aware of its relationship to the agency it has contracted to manage its DNS records. Then again, it may look like conflict because each of these two agencies claimed that implementing SPF and DMARC were the other's responsibility to initiate. In the end, though, we discovered that a lack of coordination led to the agency's non-conformance, which is another reason that a low-ambiguity, low-conflict policy may not get implemented as intended [9].

Agency conformance rates with each of the BOD 18-01 requirements are consistent with the relative risk-reward tradeoff of their implementation. Conformance was lower with requirements that, if implemented incorrectly, carry a higher risk to the domain's expected operation. Here is our justification for the relative risk-reward ratings that we assign in Table 5.5:

- Implementing SPF and DMARC incorrectly can lead to an email DoS or impersonation [4], [261], [270] and their utility depends on others' enforcement, which has not been widespread [269], [271].
- STARTTLS implementations have been fraught with vulnerabilities [262], [297] and do not guarantee any security benefit [273].
- DoS and certificate forgery are risks of incorrect HSTS implementation [298], but its benefits include mitigating MitM attacks [275].
- Disabling support for SSL, 3DES, and RC4 incurs relatively little risk, cutting off only legacy clients that cannot be upgraded [299], but avoids using these deprecated protocols and ciphersuites [276], [279], [283].

Table 5.5. A Relative Risk-Reward Assessment for Implementing BOD 18-01 Requirements, Sorted by Decreasing Conformance

Requirement	Risk	Reward
SSL	Low	High
RC4	Low	High
3DES	Low	High
SPF & DMARC	Medium	Medium
STARTTLS	High	Low
HSTS	High	High

Conformance with email security requirements was inconsistent (i.e., no apparent correlation between requirements) but was more hierarchical with Web security. While Figure 5.4 shows the conformance inconsistencies with SPF and DMARC, Figure 5.6 shows the conformance inconsistencies between SPF/DMARC and STARTTLS. The hierarchical nature of conformance with the Web security requirements is consistent with our risk-reward tradeoff assessment in Table 5.5 and is apparent in the nearly proper subsets shown in Figure 5.7.

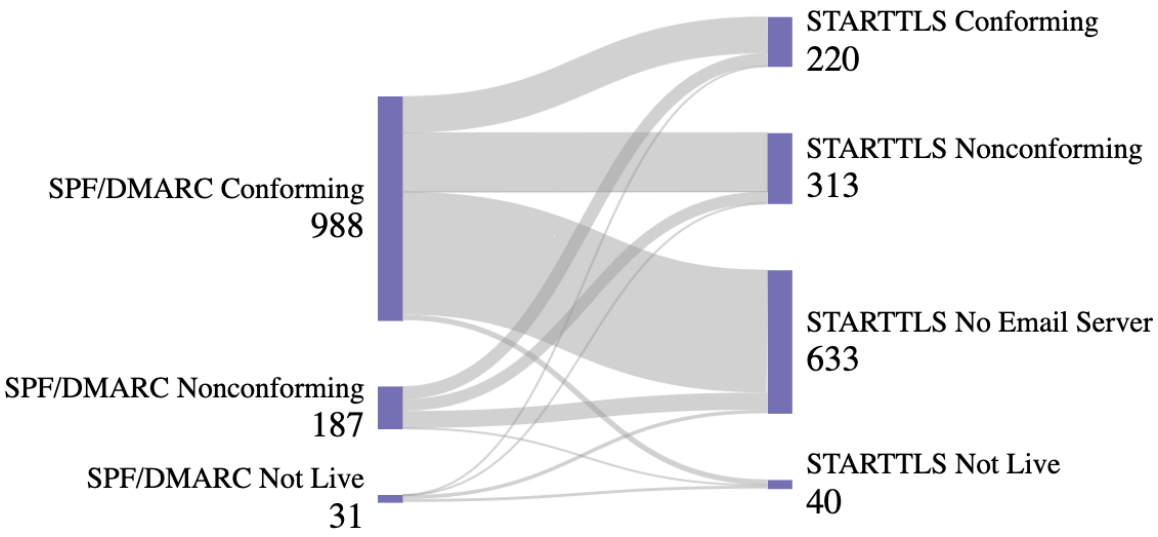


Figure 5.6. SPF/DMARC Conformance Alignment with STARTTLS Conformance. The alignment lacks any apparent correlation between the ways in which domains conform with the SPF/DMARC requirements versus the ways in which they conform with the STARTTLS requirement.

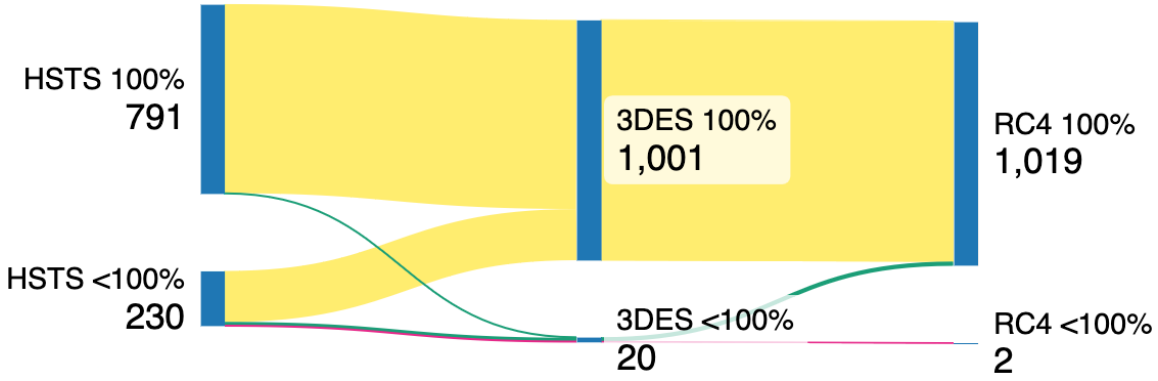


Figure 5.7. HSTS, 3DES, and RC4 Conformance Alignment. The alignment is hierarchical (trending up and to the right) in nearly proper subsets, except for three domains conforming with HSTS but not 3DES.

We focus the remainder of our analysis on SPF and DMARC conformance because of their non-binary, hierarchical configuration requirements, shown in Figure 5.1. While perusing the data, we make some interesting observations that prompted us to form the following

three hypotheses and conduct significance tests to explore factors that may contribute to non-conformance. In other words, we could not have reasonably formed these hypotheses without the measurement and RFI response data. The results of these statistical tests to determine whether significant correlations exist can help shape the direction of future research.

1. Because not all domains have MX records and our RFIs indicate confusion in how SPF and DMARC work, we first (in Section 5.3.1) hypothesized that domains without email have a higher rate of nonconforming SPF and DMARC configurations; however, we found the opposite: non-email domains have significantly higher SPF/DMARC conformance rates.
2. Second (in Section 5.3.2), because multiple domains related to offices of inspectors general (OIGs) do not conform, we hypothesized that this influential subpopulation's domains conform at a different rate than domains unrelated to OIGs; however, no significant correlation exists between conformance and whether the domain is for an OIG.
3. Third (in Section 5.3.3), because agencies that outsource their email service to a third party appeared to have higher SPF/DMARC non-conformance overall, we hypothesized that a correlation exists between email security conformance and outsourcing status; although no significant correlation exists between SPF/DMARC conformance and outsourcing, a significant correlation exists between STARTTLS conformance and outsourcing.

5.3.1 Domains without Email Service

Email spoofers can impersonate a domain regardless of whether that domain has an email service. Thus, all domains should have SPF and DMARC records to limit spoofing. Because multiple RFI responses revealed that some administrators erroneously see SPF and DMARC records as unnecessary when the domain has no email service (which we determine by the presence of an MX record), we hypothesized that domains without email have lower SPF and DMARC conformance rates. To test this hypothesis, we constructed a contingency table (cross-tabulation in Table 5.6) and used a Pearson χ^2 test between the independent variable of mail exchange (MX) record existence and the dependent variable of SPF and DMARC conformance. The Pearson χ^2 test of independence is appropriate for this hypothesis because

it measures the likelihood of variable interdependence, that is, whether one variable's distribution is contingent or conditional on another [289, p. 145]. The null hypothesis is that no correlation exists between these two variables [289, p. 147]. The assumptions of the test are as follows:

- The data are categorical [289, p. 139].
- Each domain contributes exactly one count to exactly one cell in the contingency table [289, p. 153], [291].
- The observations (measured domain configurations) are independent, as described in Section 5.2.2.
- We overcame the requirement for randomly selecting the sample by selecting all live domains from the population [291].
- For the expected values (described in Section 5.2.2), this instance meets the heuristic because the smallest expected value for Table 5.6 is 85.

We found a statistically significant correlation between domains' conformance and whether they have MX records, $\chi^2 = 16.1$ and $p = 6.07 \times 10^{-5}$, adjusted to $p = 1.82 \times 10^{-4}$ by Holm's procedure. In other words, we reject the null hypothesis that no correlation exists between these two variables and conclude, based on the expected and observed values, that domains with email service have a significantly lower SPF and DMARC conformance rate than those without email service.

Table 5.6. Cross-Tabulation of SPF/DMARC Conformance vs. MX Record Existence by Domain

	No MX Record	Has MX Record	Totals
Nonconforming	76	111	(16%) 187
Conforming	559	429	(84%) 988
Totals	(54%) 635	(46%) 540	1,175

Despite the contrary suggestion in our RFI responses, domains without email service are more likely to conform. One possibility is that SPF and DMARC configurations are much

simpler for domains without email. For example, using "`v=spf1 -all`" as a DNS `TXT` record for the domain indicates that "all" servers fail ("-") the test of whether they may transmit email for the domain [260]. Using "`v=DMARC1; p=reject`" as a DNS `TXT` record for the _dmarc.[domain] subdomain communicates to servers receiving email appearing to be from the domain that it should "reject" it [266]. On the other hand, the SPF and DMARC records of domains having complicated email configurations, with many third-party service providers and multiple internal email servers, can be lengthy and onerous to form well. Third-party email service providers include Microsoft 365 and Google Mail, plus email list distribution services like Mail Chimp and Constant Contact.

The benefit of maintaining the reputation of government-owned, non-email domains (a low-ambiguity, low-conflict policy) outweighs the cost of configuring these simple SPF and DMARC policies. For this reason, we disagree with Hu et al.'s position that for non-email domains, the cost *always* outweighs the benefit of implementing SPF and DMARC [271, pp. 98–99, Figure 4(b)]. Their argument depends on the assumption that protecting a domain from email spoofing to protect the domain owner's reputation provides "a relatively vague benefit," thus ascribing a low benefit value to reputation by default [271, p. 98]. Though they concede that "popular online services (e.g., social networks, banks)" may have sufficient incentive [271, p. 98], we (like Singanamalla et al. [300]) append online government services to this list, regardless of their popularity relative to other online services.

5.3.2 Domains for Offices of Inspectors General

Offices of inspectors general (OIGs) are structurally independent from an organization's managerial hierarchy to enable investigative work within the organization without political influence and responsible only to the organization's leader [301]. OIGs, therefore, carry great organizational authority and must maintain their integrity and unblemished reputations. Lower SPF and DMARC conformance rates in their domains enable an attacker's ability to spoof OIG email addresses, which could harm the organization and the reputation of its OIG. This is another case for a low-ambiguity, low-conflict policy-implementation.

Several nonconforming domains appear to serve OIGs. Of 21 such domains, 5 (24%) have nonconforming SPF or DMARC records, which is 150% of the live domains' non-conformance rate (16%) shown in Figure 5.4. To test the hypothesis that domains for OIGs

conform at a different rate than domains unrelated to OIGs, we constructed a contingency table (cross-tabulation in Table 5.7) and used a Pearson χ^2 test of independence between the independent variable of whether the domain is for an OIG and the dependent variable of SPF and DMARC conformance. The null hypothesis is that no correlation exists between these two variables [289, p. 147]. The assumptions of the test are as follows:

- The data are categorical [289, p. 139].
- Each domain contributes exactly one count to exactly one cell in the contingency table [289, p. 153], [291].
- The observations (measured domain configurations) are independent, as described in Section 5.2.2.
- We overcame the requirement for randomly selecting the sample by selecting all live domains from the population [291].
- For the expected values (described in Section 5.2.2), this instance meets the requirement of expected values being greater than or equal to one but not the heuristic of having expected values greater than or equal to five in at least 80% of cells. The two smallest expected values for Table 5.7 are 3.3 and 17, which means 75% of cells have expected values of at least five. We satisfy the additional assumptions of $N \geq 20$ ($N = 1,175$) and using a more conservative adjusted α of 0.025.

We did not find a statistically significant correlation between domains' conformance and whether they are an OIG domain, $\chi^2 = 0.996$ and $p = 0.318$. This being the greatest of this chapter's p values, the Holm's procedure-adjusted p value is the same [288]. In other words, we fail to reject the null hypothesis that no correlation exists between these two variables.

Table 5.7. Cross-Tabulation of SPF/DMARC Conformance vs. OIG-Related Domains

	OIG Domains	Non-OIG Domains	Totals
Nonconforming	5	182	(16%) 187
Conforming	16	972	(84%) 996
Totals	(2%) 21	(98%) 1,154	1,175

5.3.3 Domains with Outsourced Email Service

Many domains with outsourced email service do not conform with BOD 18-01's SPF and DMARC requirements. We identified domains with outsourced email providers as having a mail exchange (MX) record with a non-.gov top-level domain. Our data include 33 unique second-level non-.gov email provider domains. Grouping by provider (e.g., grouping google.com and googlemail.com together) yielded 29 providers. We hypothesized that a correlation exists between email security conformance and email outsourcing status. We tested for both SPF/DMARC and STARTTLS conformance versus outsourcing status.

To test the hypothesis that a correlation exists between SPF/DMARC conformance and a domain's or agency's email outsourcing status, we constructed a contingency table (cross-tabulation in Table 5.8) and used a Pearson χ^2 test of independence. The independent variable is whether a domain's email is outsourced, and the dependent variable is the SPF and DMARC conformance. The null hypothesis is that no correlation exists between these two variables [289, p. 147]. The assumptions of the test are as follows:

- The data are categorical [289, p. 139].
- Each domain contributes exactly one count to exactly one cell in the contingency table [289, p. 153], [291].
- The observations (measured domain configurations) are independent, as described in Section 5.2.2.
- We overcame the requirement for randomly selecting the sample by selecting all domains from the population with MX records [291].
- For the expected values (described in Section 5.2.2), this instance meets the heuristic of having expected values greater than or equal to five in at least 80% of cells. The two smallest expected values for Table 5.8 are 1.8 and 7.1, which means 83% of cells have expected values of at least five.

The difference in the conformance between domains with and without outsourced email is not statistically significant, $\chi^2 = 4.78$ and $p = 0.0915$, adjusted to $p = 0.183$ by Holm's procedure. In other words, we fail to reject the null hypothesis that no correlation exists between these two variables. Based on the expected and observed values, domains with nonconforming SPF/DMARC configurations are slightly more likely to have outsourced email than not.

Table 5.8. Cross-Tabulation of SPF/DMARC Conformance vs. Outsourced Email Status by Domain

	Not Outsourced	Mixed	Outsourced	Totals
Nonconforming	32	3	76	(19%) 104
Conforming	169	6	254	(81%) 436
Totals	(37%) 201	(2%) 9	(61%) 330	540

We find a similar perspective at the agency level for outsourced email, a slightly but insignificantly lower conformance rate for those outsourced. Figure 5.8 shows a non-linear and nonparametric correlation between the percentage of domains that an agency outsources versus the percentage of an agency's domains that conform with the SPF and DMARC requirements. These percentages are continuous ratios in which they can have any value on the continuum from 0% to 100%. To determine whether a correlation exists between agencies' domains' SPF and DMARC conformance (the dependent variable) based on whether they outsource their email service (the independent variable), we calculate Spearman's rank correlation coefficient [302]. Spearman's rank correlation is appropriate for determining the direction and magnitude of a correlation on a signed ratio scale $-1 \leq r_s \leq 1$ by assessing the degree to which the ranks of the two variables monotonically relate [303, p. 478], [304, p. 491]. The null hypothesis is that $r_s = 0$, that no monotonic relationship exists between the ranks of the two variables [304, p. 493]. The assumptions of the calculation are as follows:

- The ranked data are paired, ordinal variables [303, p. 478], [304, pp. 492–493]. We can see both these aspects in Figure 5.9.
- The data are independently randomly selected from or are representative of the population [304, p. 493]. We include all domains registered by all agencies in the population.

We ranked the percentages from low-to-high, averaged tied ranks, and computed the correlation between the ranks [303, pp. 480–481], [304, pp. 494–495]. Our calculated coefficient, $r_s = -0.0724$, indicates a slight negative correlation between outsourcing and conformance at the agency level.

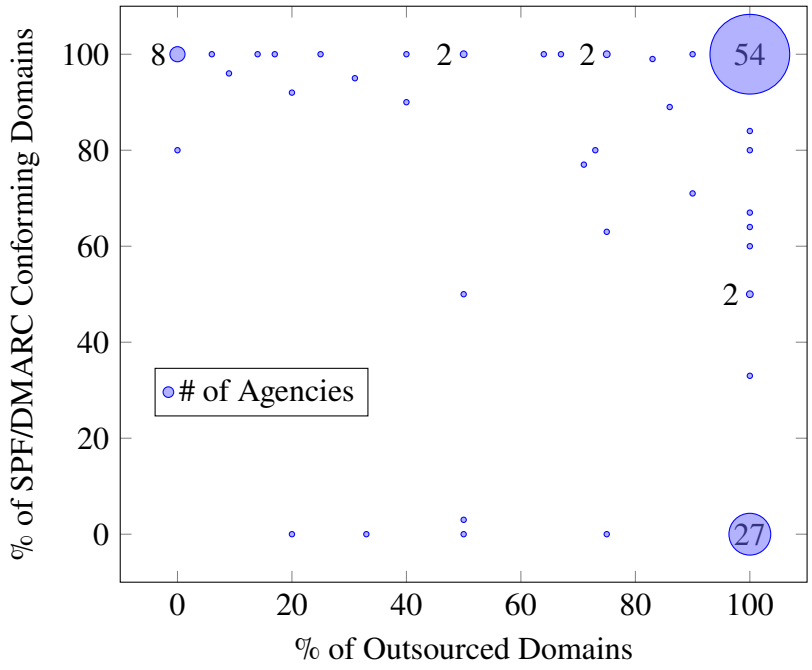


Figure 5.8. Conformance Based on Outsourcing by Agency. The bubble size indicates the number of agencies [1 – 54] (all values > 1 labeled) having that percent of domains outsourcing email service and percent of domains conforming with BOD 18-01 SPF and DMARC requirements.

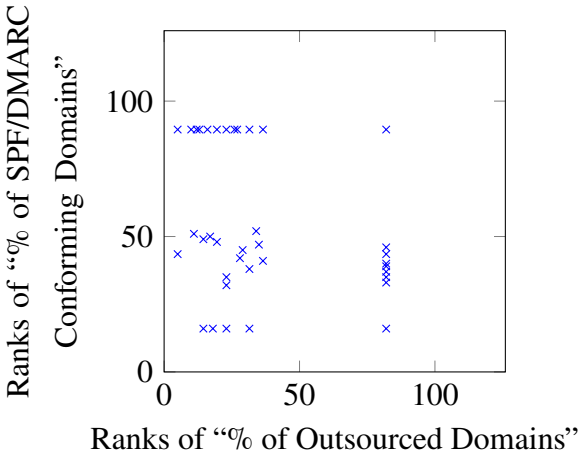


Figure 5.9. Ranked Percentages (with Tied Ranks Averaged) from Figure 5.8 of 126 Agencies' Percent of Outsourced Domains vs. Percent of SPF/DMARC Conforming Domains for Spearman's Correlation Coefficient

We tested whether outsourcing also affects STARTTLS conformance because STARTTLS is a server specific configuration. To test this hypothesis, we constructed a contingency table (cross-tabulation in Table 5.9) and used a Pearson χ^2 test of independence between the independent variable of whether a domain's email is outsourced and the dependent variable of STARTTLS conformance. The null hypothesis is that no correlation exists between these two variables [289, p. 147]. The assumptions of the test are as follows:

- The data are categorical [289, p. 139].
- Each domain contributes exactly one count to exactly one cell in the contingency table [289, p. 153], [291].
- The observations (measured domain configurations) are independent, as described in Section 5.2.2.
- We overcame the requirement for randomly selecting the sample by selecting all domains from the population with MX records [291].
- For the expected values (described in Section 5.2.2), this instance meets the requirement of expected values being greater than or equal to one but not the heuristic of having expected values greater than or equal to five in at least 80% of cells. The three smallest expected values for Table 5.9 are 3.3, 4.7, and 79, which means 67% of cells have expected values of at least five. We satisfy the additional assumptions of $N \geq 20$ ($N = 533$) and using a more conservative adjusted α of 0.025.

We found a statistically significant correlation between domains' STARTTLS conformance and whether they have outsourced email, $\chi^2 = 154$ and $p = 3.17 \times 10^{-34}$, adjusted to $p = 1.27 \times 10^{-33}$ by Holm's procedure. Because of the small p value, the suggestion to use a more conservative α, even 0.001, to reduce the risk of a type I error does not change the outcome of this instance of the test. Moreover, our "interest is in making a decision about the presence of association in the $R \times C$ table and not in estimating exact cumulative multinomial probabilities" [292, p. 1,296]. In other words, we reject the null hypothesis that no correlation exists between these two variables and conclude, based on the expected and observed values, that domains with outsourced email service have significantly higher STARTTLS conformance rates. We find this correlation magnified in the contrapositive: Of domains without outsourced email, 94% (180) do not conform with the STARTTLS requirement versus 12 domains that support STARTTLS. Higher

STARTTLS conformance among domains with outsourced email makes sense because a cloud-service provider directly influences the infrastructure's configuration.

Table 5.9. Cross-Tabulation of STARTTLS Conformance vs. Outsourced Email Status by Domain

	Not Outsourced	Mixed	Outsourced	Totals
Nonconforming STARTTLS	180	1	132	(59%) 313
Conforming STARTTLS	12	7	201	(41%) 220
Totals	(36%) 192	(2%) 8	(62%) 333	533

Note: The total here (533) differs from the totals in Tables 5.6 and 5.8 (540) because some agencies changed email servers between our July DNS queries and our November STARTTLS probes.

5.3.4 More on Outsourcing and Email Cloud-Service Providers

We mentioned some effects of outsourcing U.S. Federal Executive Branch agencies' email services to cloud-service providers in Sections 5.2.1 and 5.3.3, including reporting an insignificant difference in the SPF/DMARC conformance between domains with outsourced email and those with agency-hosted email services. Hu et al. reported that email administrators find configuring DMARC to be a greater burden when email is outsourced to a cloud-service provider, and are therefore less likely to implement them [271]. However, email administrators should find the opposite, that email cloud-service providers, which specialize in provisioning email services, would encourage and make configuring SPF and DMARC easier. To support this vision, we consider the most commonly occurring misconfigurations from our data against the documentation of the email cloud-service provider most used in our data. This analysis enables us to form three recommendations for email cloud-service providers:

1. Offer a default DMARC record.
2. Remind domain owners to upgrade to reject policies.
3. Improve documentation for report-receiver records.

When organizations have contractual relations with third parties to implement policy, it does not relieve implementation responsibility from the organization tasked by the policymaker. This is characteristic of microimplementation across a diverse enterprise, in that organizational components have autonomy to implement a policy how they see best. Twenty-eight of the 126 U.S. Federal Executive Branch agencies with registered .gov domains maintain non-outsourced email servers while also subscribing to outsourced email providers. These 28 agencies have a widely varying proportion of domains with outsourced email (range [6% – 90%], median 50%, sample mean 51%, and sample standard deviation 25.9%). Within some agencies, component organizations manage their own Web and email instead of the agency managing these capabilities centrally, which could entail multiple email providers. Differing missions and functions may dictate this separation. When delegating service provisioning, agencies must serve as an intermediary in the policy hierarchy (see Figure 1.2) so that lower levels can comply with policies. From this widespread variation in microimplementation, we can learn more about which implementation methods are most effective at reducing the policy-implementation gap.

The email cloud-service provider with the greatest use by U.S. Federal Executive .gov domains is Microsoft. Of 87 agencies with 196 domains outsourced to Microsoft for email, 33 (38% of 87) agencies have 58 (30% of 196) nonconforming domains. The relative order of these domains' reasons for non-conformance (Table 5.10) is like the order of reasons for the population of live domains (Table 5.1).

Table 5.10. How Domains with Microsoft-Outsourced Email Fail to Conform with SPF and DMARC Requirements

Nonconforming Requirement(s)		***n***	**%**
No DMARC Record	(Total 24)	23	39.7
Reject Policy Not Set and NCCIC not in RUA		11	19.0
Invalid DMARC Record Syntax	(Total 10)	8	13.8
Reject Policy Not Set	(Total 19)	4	6.9
NCCIC not in RUA	(Total 17)	4	6.9
No RUA or RUF and Reject Policy Not Set	(Total 3)	3	5.2
Invalid SPF Record Syntax	(Total 2)	2	3.4
Invalid DMARC Record Syntax, Reject Policy Not Set, and NCCIC not in RUA		1	1.7
Invalid DMARC Record Syntax and NCCIC not in RUA		1	1.7
No SPF Record and No DMARC Record	(Total 1)	1	1.7
Total		**58**	**100**

Recommendation #1: Offer a default DMARC record. Not having a DMARC record is the most common nonconforming requirement in the population and in the subset of domains that outsources email to Microsoft. Having a default DMARC record could prove helpful [305], a concept that Simoiu applies to cloud-hosted Web server security [306]. Like the no-email-service default DMARC record that we offered in Section 5.3.1, government-contracted email cloud-service providers could use this starting configuration for agencies to customize:

```
v=DMARC1; p=quarantine; rua=mailto:reports@dmarc.cyber.dhs.gov;
```

This syntactically correct DMARC record complies with BOD 18-01's DMARC requirements, except for upgrading from a quarantine policy to a reject policy [19]. Furthermore, it is consistent with one of Microsoft's example DMARC records [307]. For non-U.S. Federal government agencies, a cloud-service provider could simply personalize the default by

pre-filling the RUA with an organization's administrator's email address. The quarantine policy strikes a balance between causing a DoS and allowing all email despite SPF or DKIM failing [266].

Domains with email outsourced to Microsoft nearly always conform with the SPF requirements (98.5%, 193 of 196), unlike the sample of live domains (92.9%, 1,092 of 1,175, shown in Figure 5.4). A deeper analysis of the process to establish or transfer to a cloud-service provider-hosted email service may reveal why SPF conformance is so high compared to DMARC conformance. Interestingly, one domain's lack of an SPF or DMARC record contradicts Microsoft's SPF documentation. Having outsourced its email to Microsoft with a custom domain (i.e., a .gov), "Office 365 requires that you add a Sender Policy Framework (SPF) `TXT` record" [308]. We lack an RFI response from this domain's agency.

Recommendation #2: Remind domain owners to upgrade to "reject" policies. Not setting a reject policy is likewise the second most common nonconforming requirement. These domains have either a "none" (12 of 196, 6.1%) or "quarantine" (7, 3.6%) policy. Like BOD 18-01, Microsoft documentation recommends setting a "reject" policy after administrators have reviewed failure and aggregate reports (described in Section 5.1.1) at a "none" or "quarantine" policy over a sufficient period [19], [307]. Although we have no evidence pointing at a reason for not escalating the policy over time, forgetfulness and being preoccupied with other cybersecurity management responsibilities are among the possibilities [147]. Hureau et al.'s longitudinal DMARC study reflects some DMARC policy escalation over time but shows that nearly all domains maintained their policy levels [270]. The results of recent studies on notification indicate that CISA and email cloud-service providers sending periodic DMARC policy upgrade reminders to agencies and clients not having reject policies may help achieving a greater implementation rate [294], [295]. To verify this, we recommend a longitudinal study in which a test group receives notifications reminding domain owners to upgrade their DMARC policies, complemented by surveys or interviews to discover the reason for having "none" or "quarantine" policies instead of "reject."

Administrators also have not been including the NCCIC address in the RUA, the third most nonconforming configuration. Enabling CISA's ability to analyze DMARC aggregate

reports across the .gov enterprise can inform future actions. As with setting a reject policy, periodic reminders to owners of domains without the NCCIC address in the RUA may improve this relatively simple requirement.

Recommendation #3: Improve documentation for report-receiver records. Having a DMARC record with valid syntax is the fourth most missed requirement for both the live domains and this subset (10, 17%). Although CISA describes this requirement in the implementation guidance on the BOD 18-01 Web page [19], the Microsoft documentation omits mention of report-receiver records [307], despite their need for these external RUAs and RUFs [266]. We therefore recommend email cloud-service providers describe the necessity for report-receiver records.

The fifth and sixth most missed requirements are not having any RUA or RUF and having a syntactically invalid SPF record. Unlike the explicit direction in BOD 18-01 for configuring an RUA, the Microsoft DMARC guidance only implies that servers receiving email need the originating domain's RUA to send reports generated about email from that domain [307]. From Table 5.10, the two invalid SPF syntaxes are neglecting the field delimiter and too many DNS lookups. Although BOD 18-01 does not mention either of these, the Microsoft SPF documentation mentions but does not enforce them [308].

Our three recommendations to service providers are consistent with those of Czybik et al. [264], for example, in providing documentation to inform customers of the benefits of SPF and DMARC and how to configure them. Because cloud-service providers control their infrastructure's configuration, they can leverage defaults in its configuration. Still, as just explained, customers might not use the default. Future research could study email cloud-service providers' assistance and encouragement for SPF, DMARC, and STARTTLS versus the ratio of their clients that conform.

5.4 Observations and Limitations

Through our exploration of the factors contributing to this policy-implementation gap, we discover the need to further explore how U.S. Federal Executive Branch organizations operationalize their cybersecurity policies. The data show that we cannot predict policy conformance by characteristics like ease of configuration, email outsourcing, or resourcing levels. This suggests that we lack the correct features with which to model conformance.

Within the ambiguity-conflict model, we categorize BOD 18-01 as a low-ambiguity, low-conflict policy [9]. We should find it implemented wherever directed organizations have sufficient resources to do so.

Although some statistics reported in this chapter indicate strong policy implementation, BOD 18-01 is inadequate to achieve state-of-the art cybersecurity protections. For example, it does not require state-of-the-art configurations, like deprecating TLS 1.1 and International Data Encryption Algorithm (IDEA) ciphersuites [18]. Only 800 (78% of 1,021) domains support only TLS 1.2 – 1.3, which is at least 13% better than the average government Web server across the globe [309]. Treating policy as a minimum performance standard for security misses key vulnerabilities.

Responses to our requests for information (RFIs) may have been limited by the notion that responding to a stranger's unsolicited RFI could discredit the agency, an individual, or cybersecurity team. The RFI asks why the agency is delinquent in conforming with a mandatory policy. Answering the RFI first requires admitting the shortfall exists to an unknown outsider claiming to have measured it, an action for which people have differing propensities [310]. Our efforts to build trust to overcome this limitation include using our DoD credentials for identification, applying a Secure/Multipurpose Internet Mail Extensions (S/MIME) cryptographic signature to each email conveying an RFI, and promising the confidentiality described in Appendix C. Although Stock et al.'s experience indicates that digital signatures do not help [294], our institution (as another Federal entity) may be better-known to our recipients who also can verify our digital signatures through the DoD public key infrastructure (PKI). No agency used the authors' controlled unclassified information (CUI)-accredited repository for submitting a response and we lack any indication of how many of the 10 responding agencies responded because of the authors' affiliation. Thus, we believe this study is repeatable by other U.S. researchers.

5.4.1 Conformance versus Complexity

We counter-intuitively find that configuration complexity and conformance correlate (see Section 5.3.1). Although U.S. Federal Executive Branch .gov domains without email have simpler SPF and DMARC configurations, administrators often fail to protect the domain from becoming a spoofed source of email. Simultaneously, we recognize in Section 5.1.1 that

greater configuration complexity increases others' burden to help and oversee cybersecurity configurations. Independent entities provide domain registration and host DNS records so enforcement is impractical during registration. Even though CISA provides simple SPF and DMARC records on the BOD 18-01 Web page [19] for cases of not having an email server, like we offer in Section 5.3.1, some agencies still do not implement them.

Responses to our RFIs often indicated a lack of understanding among respondents of how SPF and DMARC help protect a domain from impersonation. Although BODs are likely not the right medium by which to provide background information on how specific protocols function, CISA's follow-up on the BOD 18-01 Web page [19] provides this information. What additional ambiguity-reducing resources should the policymaker provide for the practitioners whose confident misunderstanding of the policy has prevented that policy's implementation?

Conformance reminders may help system administrators who forget to conform [147]. Because some domains' conformance increased after our RFIs (see Section 5.2.4), as in other SPF/DMARC studies with notifications to domain owners [264], [269], [270], some cybersecurity teams at the affected agencies could have forgotten about BOD 18-01 and receiving an RFI prompted them to action. We know this was the case for the two responses (described at the end of Section 5.2.3) that attribute their conformance to our RFI; however, we acknowledge that our lack of a control group and relatively small sample size prevents us from claiming a causal relationship.

We do not know how often the DHS or CISA reminds agencies of their compliance obligations. We described in Section 5.3.3 the value of an intentionally slow, step-wise implementation of escalating cybersecurity policy, like progressing from a DMARC policy of "none" to "quarantine" to "reject." However, reminders may be an annoyance when forgetfulness is not to blame. Then organizations must adopt other policy strategies.

5.4.2 Semantics

Hidden Semantics

A policy-implementation gap occurred in some U.S. Federal Executive Branch agencies between the semantic goal stipulated in BOD 18-01 and the syntactic policy validation

that `trustymail` performs (discussed in Section 5.1.1). Without our customization of the `trustymail` analysis to include the semantics of the selected qualifier on the `all` mechanism, we found conforming SPF records that do not provide the intended implementation. This finding also furthers the problem with checklist compliance: that the mere completion of such a checklist does not necessarily secure the system [91].

SPF's relative simplicity can become complex for organizations with many email servers sending on its behalf. For example, one RFI response describes:

> Our employees and their collaborators create lists at effectively all the major email service provides and even other organizations. Specifying a reject policy would create a situation where we would block the ability (the list source is not in SPF) for this sort of collaboration to occur. We have no desire to limit out [*sic*] collaborators to specific email list providers.

If the SPF record does not list all these services' servers, email from them may be quarantined or rejected by receiving servers. The organization-specific knowledge required to check for semantic correctness in complex contexts like this makes superintending these records intractable for federated IT oversight. Pushing IT functions lower in an organization's structure, however, makes compliance uniformity harder to achieve. Agencies must find the right balance between centralization and decentralization of cybersecurity management.

Registration Semantics

The survey data revealed a few instances of microimplementation conflict, primarily in questions of authority. All the domains in this study are registered as "Federal – Executive," for which there are 126 owning agencies. Legally, CISA's authority to issue and enforce BODs affects all Federal Executive Branch agencies, except for National Security Systems (NSSs) and mission-critical DoD and Intelligence Community (IC) systems [254]. CISA claims authority over 97 of these legally authorized agencies (see Section 5.1) but not any DoD or IC agencies [259]. Whether a public .gov domain is an NSS or is mission-critical is not always readily apparent. In the absence of insider knowledge of the purpose of each of these domains, we do not attempt this determination, but raise the question of whether the apparent blanket exception of these agencies' .gov resources helps the .gov enterprise's cybersecurity.

Of the agencies with "Federal – Executive" registrations, some (fewer than ten, each with a single registered domain) may fit better in another category. For example, the "Interstate" category may be a better fit for the Delta Regional Authority (DRA). We do not know whether this option was available when the DRA requested a .gov domain and if not, whether anyone at the .gov Registrar reviewed the existing list upon creation of the "Interstate" category to see what organizations (like the DRA) should re-designate into it. Notwithstanding, we accept their "Federal – Executive" registrations for this study.

Agencies register second-level domain names with the .gov top-level domain Registry but manage their own subdomains. We queried only second-level .gov domains for SPF and DMARC conformance. Because an organization's subdomains can have separate SPF or DMARC records, many more additional subdomains could be included for measuring government agencies' BOD 18-01 conformance [260], [266]. Because additional non-trivial work is involved in verifying subdomains and things unchecked have a greater likelihood of being in an undesirable state [93], our measurements are likely upper bounds of conformance rates.

5.5 Conclusion

Organizational factors unrelated to security hindered cybersecurity policy conformance in this study of policy implementation across the U.S. Federal Executive Branch. We use the term *policy-implementation gap* to describe the dichotomy between the cybersecurity posture that an organization's management intends to effect and that which it finds as an outcome to the policy's implementation [10]. We analyze the policy-implementation gap through Matland's ambiguity-conflict model of policy implementation [9]. We also adapt the policy-implementation gap measurement model (Equation 1.2), devising a distance-to-conformance model (Equation 5.1, described in Section 5.1.3) to measure a configuration's degree of policy implementation.

What can cause the implementations of a mandatory cybersecurity policy to vary across U.S. Federal Executive Branch organizations? As a top-down policy from a government agency to agencies across the U.S. Federal Executive Branch, BOD 18-01 is a low-conflict policy, and its precise requirements lead to its low ambiguity. The low-ambiguity, low-conflict combination indicates that once an agency can resource the implementation that it

will implement the policy (see Figure 2.1). Our query and server probing data corroborate that most agencies have implemented most policies on most of their domains. However, in several agencies, confusion about how the directed requirements work has caused policy ambiguity, and a few agencies have experienced conflict in implementing this policy. In these cases, the ambiguity and conflict have occurred at the microimplementation level. The discovery of these special cases enriches the study of the policy-implementation gap and the value of using the ambiguity-conflict model as a lens for analyzing policy non-conformance.

We focused on SPF and DMARC conformance because of their non-binary and hierarchical configuration requirements. To enrich the collected DNS query and server probing data with reasons for agencies' non-conformance, we notified and surveyed the nonconforming agencies by sending each a request for information (RFI). Among the responses, we find a lack of understanding of how SPF and DMARC function, misunderstandings of and conflicts with authorities and responsibilities, insufficient resources, and operational reasons for nonconforming. After having allotted 90 days for reacting to our RFIs/notifications, we re-queried their DNS records to measure their effect, finding significant conformance improvement across second-level .gov domains. All but one of the improving domains were subject to our RFIs.

Our findings suggest aspects of configuration complexity, organizational structure, tracking and incentivizing conformance, affordances in how agencies register domains, and syntactic versus semantic accuracy. We observed that these agencies' understanding of how each cybersecurity policy interacts within the context of their information systems, and how each impacts cybersecurity, affects the policy's effective implementation.

CHAPTER 6:
Discussion and Conclusion

To decrease the frequency and severity of cybersecurity policy-implementation gaps in organizations, we identify contributing factors across diverse cybersecurity studies. We believe that reducing these gaps can lead to fewer and less-severe incidents than those cited at the beginning of Chapter 1, including data breaches [2] and businesses losing billions of dollars because of misconfigured cloud-based services [5]. This was the motivation for our primary research question, "What non-adversarial sources can cause inconsistencies in implementation and the realized outcomes that differ from a cybersecurity policymaker's stated intentions?"

We identified many factors contributing to policy-implementation gaps in cybersecurity in the previous chapters. Typically, insufficiencies in implementation can be caused by policymakers appearing to misunderstand the nuances of implementation and practitioners appearing to misunderstand the reasons for a policy. Both misunderstandings are typically provoked by human and organizational factors. This dissertation has five primary contributions:

1. Using the ambiguity-conflict model for analyzing the policy-implementation gap and, for the first time, for organizational cybersecurity policy, our diverse studies broaden the model's usefulness and offer an alternative perspective for understanding the policy-implementation gap in cybersecurity.
2. We conducted two studies on information technology (IT) capability nonuse not yet addressed by the Navy by surveying administrators and integrating technology-use mediation concepts with the ambiguity-conflict model.
3. We examined the conditions for privacy-risk perception and mitigation in real privacy incidents against a standard privacy framework, publishing an analysis of how the National Institute of Standards and Technology (NIST) Privacy Framework could have more adequately helped to identify the risks exploited in the studied incidents.

4. By gathering data on both network-services and information-security workers, we explored how the policy-implementation gap manifests across the U.S. Federal Executive Branch when implementing a legally mandated cybersecurity policy.
5. We devised a model for organizations to measure the policy-implementation gap and demonstrated how to use the model for measuring the gap across an enterprise cybersecurity policy's implementations in the U.S. Government (USG).

This dissertation contributes to the field of organizational cybersecurity, whose members study the "efforts organizations take to protect and defend their information assets, regardless of the form in which those assets exist, from threats internal and external to the organization" [14, p. 4]. The field integrates knowledge from computer and social sciences and offers a more holistic perspective of the policy-implementation gap than is usually considered [10]. Our work contributes to the study of organizational cybersecurity by exploring shortfalls in cybersecurity policy implementation in organizations.

6.1 Research Studies

Chapters 3, 4, and 5 explored three aspects of organizational cybersecurity within organizations' adoption and use of software systems within the information-systems life cycle [13]: the adoption of new technologies, using standards and frameworks, and enterprise cybersecurity in the USG. For each aspect, we offered a research question and listed our contributions.

In each study, we saw how the ambiguity-conflict model of policy implementation explained the circumstances that contributed to the observed policy-implementation gap [9]. As the separation (especially in organizational structure) between cybersecurity policymakers and practitioners implementing the policy increases, so does the opportunity for a larger policy-implementation gap. Even with adequate resources, the non-functional goals of resilience, privacy, and cybersecurity can fail due to their dependence on the system, user, cultural, and structural interactions. We can adapt Figure 2.1 to show how our studies fit in an ambiguity-conflict matrix in Figure 6.1 from both the macroimplementation and microimplementation perspectives. Although Matland reconciles these perspectives in a unified model for public-policy implementation, our studies showed that the levels of ambiguity and conflict in organizational cybersecurity policy differ between the macroimplementation

and microimplementation perspectives. From the macroimplementation perspective, the policies had low ambiguity and low conflict. On the other hand, the perspective from the microimplementation level showed greater ambiguity and conflict. Enterprise cybersecurity in the USG, studied in Chapter 5, has mostly low ambiguity and low conflict, and its implementation is mostly a matter of resources. Both the chat afloat study in Chapter 3 and the standards-and-frameworks study in Chapter 4 have mostly high ambiguity and low conflict, for which we hypothesize, consistently with the model, that implementation would be successful when sufficient autonomy exists at the microimplementation level. The communications shifts study in Chapter 3 represents a high-ambiguity and high-conflict policy whose implementation occurs upon forming a coalition at the microimplementation level.

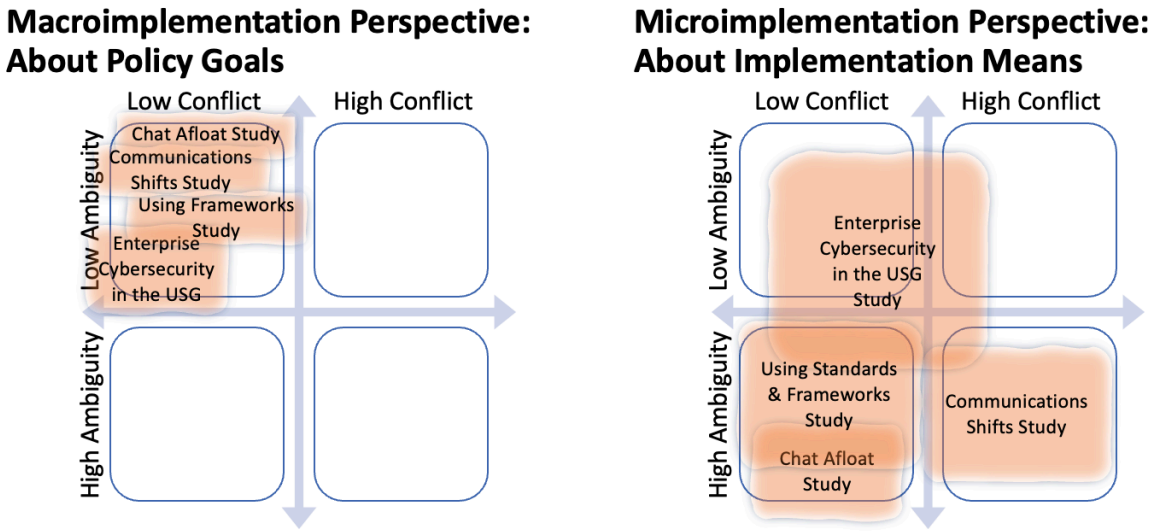


Figure 6.1. How Our Studies Overlay the Ambiguity-Conflict Matrix at the Macroimplementation and Microimplementation Levels. Adapted from Figure 2.1.

In exploring the research question, "How can we assess the magnitude of a policy-implementation gap?," we considered the appropriateness of a quantified model in each study. Based on the concepts described in Section 1.1 and the model offered by Equation 1.2, our studies vary as to how well a quantified model explained the gap. The study for which a measurement is least appropriate is the communications shift study in Chapter 3 because factors affecting policy implementation were inconsistent in the unique cases that our survey respondents

described. Secondly, uncertainties in the analyzed privacy-inference incidents hindered their accurate quantification (see Table 4.2). Quantifications with too much uncertainty can mislead and engender a false confidence due to their apparent precision [225]. Our study on chat afloat could use the model offered in Equation 1.2 (demonstrated in Section 3.2.4), because the policy and its implementation are straightforwardly comparable. As described in Section 5.1.3, Binding Operational Directive (BOD) 18-01's explicit policy requirement-to-control mapping provided the best opportunity of our studies to measure conformance. We used the model described in Section 1.1 to devise a distance-to-conformance model (described in Section 5.1.3) by modeling the hierarchical policy requirements as a tree and summing their values using Equation 5.1.

6.1.1 IT Adoption

In Chapter 3, "Adopting Resilient Information Technology Afloat," we asked, "How can the level of adoption of available IT capabilities contribute to gaps between policy and practice?" We conducted two studies on sailors' use of the U.S. Navy's afloat tactical networks, one on the resilience of Internet Protocol (IP)-conveyed command and control (C2) communications over chat, and one on communications shifts during cross-area of responsibility (AOR) changes of operational control (CHOPs). These studies involved technical configuration analysis and surveys of sailors in command, control, communications, and computers (C4) leadership positions responsible for the provisioning and use of these systems. We found that these systems were not fully "adopted;" that is, their acquisition, implementation, and initial user training were incomplete [13]. Because of this, their operation in the continuance phase of the information-system life cycle revealed a policy-implementation gap. We offer the following insights, expanding on contribution 2 above:

1. An organization's demanding work environment, limited training on resilience-providing capabilities, and system complexities can overpower its members' ability to use their discretionary professional judgment in implementing the full technical capabilities in information systems.
2. Most surveyed C4 leadership do not use resilience-providing capabilities of applications in afloat tactical networks because they are unaware of these capabilities and their benefits.

3. Factors external to an organization can impact the organization's implementation actions. Surveyed ships routinely receive insufficient support of their communications shifts during cross-AOR CHOPs, often resulting in implementation conflict and unplanned outages. Although some C4 leaders in the fleet are aware of this, our survey responses from well-provisioned ships (those with multiple satellite communications (SATCOM) paths) document a sample of the situation.
4. The Navy's acquisition, fielding, and nonuse of specific communications architectures could be more resilient by design with some relatively simple technical changes that the Navy has not yet pursued.
5. The ambiguity-conflict model [9] and technology-use mediation concepts [128] helped explain the survey respondents' underuse of resilience-providing capabilities and justify our intervention recommendations. High ambiguity among a policy's implementers leads to diverse microimplementations that worsen the policy-implementation gap. These phenomena decouple the systems' sponsors' policy requirements from their implementations and increase ambiguity in the deployed systems' use. In some cases, conflict for limited resources in implementation increases the gap.

Our research from Chapter 3 suggests the following recommendations:

1. Improve these systems' organizational metastructuring: Requirements owners and resource sponsors can champion the capabilities of deployed systems to empower C4 leadership to serve as technology-use mediators, and programs of record (PORs) can design systems to use resilience-providing capabilities by default.
2. Alternative Navy communications system design choices for applications to support C2 communications include:
   (a) Enabling and scheduling periodic Domain Name System (DNS) zone transfers from fleet network operations centers (NOCs) to ships.
   (b) Establishing static Exchange send-connectors to enable email delivery directly between ships without depending on a NOC ashore.
   (c) Establishing procedures and standardized configurations for operating a distributed chat architecture and designing future chat systems to prefer resilience-providing capabilities by default.

3. Future research should focus on three aspects that our data suggest will routinize these shifts:
   (a) Streamlining the satellite access request (SAR) process and satellite access authorization (SAA) execution.
   (b) Simplifying the wide-area network (WAN) architecture that currently limits load balancing and prioritization of IP traffic classes across geographic boundaries.
   (c) Informing sailors serving afloat and at NOCs ashore of the capabilities, limitations, and operating configurations of afloat tactical-network components.
4. A notional optimized communications-shift procedure for ships (like that which we offered in Supplemental 2) is needed for when they have multiple SATCOM missions to shift.

6.1.2 Using Standards and Frameworks

Chapter 4 addresses the question, "How can standards or frameworks guide or hinder organizations' policy implementation?" We analyzed four privacy incidents, caused by inferences, against the NIST Privacy Framework [43], assessing the degree to which the framework could have helped to identify the risks associated with these incidents. This analysis revealed limitations in the framework's ability to help organizations identify inference risk. Moreover, when organizations use the framework as a checklist (against NIST's instructions), their success depends on the professional judgment of privacy and cybersecurity practitioners who are better able to recognize its shortfalls than anyone blindly implementing mapped controls. In other words, every adopting organization requires a unique application of the framework. We offer the following insights, expanding on contribution 3 above:

1. Organizations' privacy-policy implementation shortfalls included failing to account for an attacker's ability to combine anonymous data with public information, relying on the removal of personally identifiable information (PII)-category data to protect people from reidentification, and having difficulty translating policy into a technical implementation.
2. Standards and frameworks can help organize work practices and guide professional judgment but are insufficient as independent aids. Organizations must go beyond a framework's explicit specifications to infer intended goals, customizing their approaches to the organization's unique needs.

3. We corroborated previous results: Blind reliance on frameworks (i.e., checklist compliance) is insufficient. If organizations comply verbatim with the NIST Privacy Framework, their privacy-management programs may not achieve their goal of protecting privacy. Instead, every organization requires a unique application of the framework to account for its unique needs, guided by professional judgment.
4. Because of the framework's by-design ambiguity, for organizations to implement the NIST Privacy Framework according to their specific circumstances, an adequate implementation requires sufficient autonomy at the microimplementation level in each adopting organization [9].

We discussed our findings with the Privacy Framework group at NIST and submitted feedback for incorporating into the framework's mapped controls in NIST Special Publication (SP) 800-53r5 [46]. We also offered grounded recommendations to help organizations overcome the shortcomings faced by the organizations in our analyses:

1. Map existing controls related to inferences to the inferences subcategory in the framework.
2. Increase awareness and best practices within the organization related to inferences by training organizational members about inferences.
3. Update controls pertaining exclusively to PII such that they adequately address inferences.
4. Improve the translation of policy into technical requirements by positioning people who can translate between policies and implementations in policy-making and policy-implementing positions.

6.1.3 Enterprise Cybersecurity in the U.S. Government

For Chapter 5, we asked, "What can cause the implementations of a mandatory cybersecurity policy to vary across U.S. Federal Executive Branch organizations?" We analyzed the requirements of Department of Homeland Security (DHS) BOD 18-01 [19] and U.S. Federal Executive Branch agencies' conformance to these requirements on their .gov domains. It is a top-down low-conflict policy, and its precise requirements provide low ambiguity, enabling its implementation as soon as sufficient resources are available. Our analysis involves querying and probing servers to measure conformance, surveying non-conforming agencies,

and re-querying servers to measure changes in conformance after agencies received our requests for information (RFIs). Most agencies have implemented most policies on most of their domains. However, in several agencies, confusion about how the directed requirements operate has caused some policy ambiguity, and a few agencies have experienced some conflict in implementing this policy. The discovery of these special cases enriches the study of the policy-implementation gap and the utility of using the ambiguity-conflict model as a lens for analyzing policy non-conformance. We offer the following insights, expanding on contribution 4 above:

1. We used the concept of measuring an implementation's conformance to a policy (introduced in Section 1.1) to offer a *distance to conformance* model.
2. When agency components have their own information-system and cybersecurity governance, the heterogeneity of the organization's cybersecurity policy implementation can be similarly diverse, despite having a common policy.
3. Reasons for agencies' non-conformance include misunderstanding the cybersecurity controls specified in the policy, resource priorities, forgetfulness, questions of authority and responsibility, and operational conflicts.
4. As the uniqueness of agencies' needs increased, so did the oversight complexity, as each can raise a reason for deviating from a detailed one-size-fits-all enterprise-wide policy.
   (a) Greater email configuration complexity in agencies' domains reduces a higher headquarters' ability to meaningfully help and oversee cybersecurity configurations. An overseer using `trustymail` can verify the syntactic correctness of implemented policy but may lack the component-specific knowledge to verify the semantic correctness.
   (b) Policy directing specific configurations becomes untenable to enforce as agencies and their components' diverse needs prompt exceptions to the policy.
   (c) Syntactic conformance does not imply semantic conformance.
   (d) Inefficiencies in IT provisioning and management arise in the USG, resulting in disparate configurations, including for cybersecurity-policy implementations.
5. Even with a low-ambiguity and low-conflict policy, organizational factors can introduce ambiguity and conflict that hinder implementation at lower levels in the USG.
6. Treating policy as a performance standard for security can miss key vulnerabilities.

7. Mandating detailed configurations does not necessarily yield stronger security.
8. A lack of understanding of how the directed cybersecurity policies function negatively impacted agencies' conformance.
9. Researcher notifications to agencies with nonconforming domains correlated with improved conformance.

Our analysis in Chapter 5 enabled us to form three recommendations for email cloud-service providers about their U.S. Federal Executive Branch clients' Domain-based Message Authentication, Reporting, and Conformance (DMARC) record configurations:

1. Offer a default DMARC record.
2. Remind domain owners to upgrade to reject policies.
3. Improve documentation for report-receiver records.

We discover the need to further explore how organizations operationalize their cybersecurity policies. Qualitative studies with implementers would best serve this end.

6.2 Opportunities for Future Research

Each of our studies presents opportunities for future research. An overarching question arises about policy implementation in organizations: Can better policy design choices enable better implementations? We suspect not for the reason implied in this question: If such a formula for better policy design exists, how would it differ across the three different forms of policy-implementation gap and different types of organizations presented in this dissertation? In other words, differences abound in cybersecurity-policy implementation situations. Despite researchers over many years studying procedural compliance, usability of cybersecurity capabilities, cybersecurity training and education, and requirements formalization, cybersecurity policy implementation remains a challenge.

From our IT adoption studies, future research could explore how best to employ tactical communications networks afloat, combining ethnography, system and protocol configurations, and field studies. Why has the Navy not previously identified or, if identified, not acted on our proposed system design interventions, especially those that are relatively simple to implement? We also propose investigating improvements to the military satellite-access request, authorization, and execution processes, and an analysis of the effectiveness of

the Department of Defense (DoD)'s Adaptive Acquisition Framework's annual-assessment mechanism [144]. We discovered these factors about design and processes during our study on the nonuse of resilience-providing capabilities and hypothesize that by acting on their answers, the Navy could expand these systems' use to improve communications resilience.

From our study on standards and frameworks, we propose further analysis of them, combining validation of requirements with controls, verification of control implementation, and ethnographic work to gauge and improve policymakers' and practitioners' level of understanding. This research work would be especially insightful if conducted by organizations implementing these standards and frameworks. Starting from our study, at least two other directions are available to pursue. One, with its focus on cybersecurity compliance, the USG would likely be interested to know how the practice of using frameworks as checklists, and then demanding organizational components document deviations from them, help an organization's cybersecurity posture. Two, specific to privacy inferences, more work could expand our taxonomy of inference types to formalize the types and provide more examples of each type. A more formal analysis of inferences could lead to their more-common inclusion as considerations in organizations' privacy programs.

From our study of enterprise cybersecurity in the USG, an ethnographic approach to conduct interviews and field studies in nonconforming agencies could improve our understanding of policy implementation across large and complex organizations. Such an effort would require many years of work. In addition, we propose studying how email cloud-service providers' assistance and encouragement to their clients to implement best-practice cybersecurity policies could affect clients' adoption of these cybersecurity controls.

Future work can develop our gap measurement model in multiple directions beyond the scope of the studies in this dissertation. First, we could use it in other scenarios that involve measuring an implementation's conformance to a policy. A next step is to try using the model in other observational studies to validate its value and discover additional shortcomings that the studies in this dissertation did not expose. Other developments that would increase its complexity would take from its simplicity, raise the learning curve to use it, and incur additional costs for organizations. For example, one could expand the model to account for the implementation costs (time, software tools, hardware upgrades, etc.) of each policy

requirement. A goal of this future work could be to develop metrics that help organizations' cybersecurity efforts.

Having used Matland's ambiguity-conflict model of policy implementation for analyzing the policy-implementation gap in organizational cybersecurity policy, others could also use it for finding contributing factors to cybersecurity policy implementation failure. As a political scientist and publishing it in the Journal of Public Administration Research and Theory, Matland's model naturally applies to government policies and government organizations [9], which are the focus of this dissertation. To show how the type of organization can also affect policy implementation [69], one could find a cybersecurity policy that applies to the U.S. Navy and non-Navy USG organizations, conduct a similar analysis to that presented in Chapter 5, and compare the findings between these two organization types. We could apply the ambiguity-conflict model to organizational policies unrelated to public policy or governments at any level. Further research can continue exploring unique cases to broaden and deepen our understanding of the policy-implementation gap in cybersecurity.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A:
Afloat Chat Server Survey

We received Institutional Review Board (IRB) approval to conduct this survey and Navy Survey Office authorization with control number NSPM23.14, expiring June 14, 2025.

We solicited survey participants through the Navy Information Professional Officer's milSuite page [132] with cross-posts into Information Professional Officer Teams channels on Flank Speed [133] and at the Naval Postgraduate School (NPS). The solicitation advertised that the survey supports research on improving the resilience of command and control (C2) communications in afloat tactical networks for the purpose of helping the Navy overcome sociotechnical barriers to using capabilities inherent within fielded systems. Most participants took between 5 and 15 minutes to complete the survey. Sailors could take this survey once for each Consolidated Afloat Network and Enterprise Services (CANES) ship in which they served in an information technology (IT) leadership position.

The questions are as follows:

1. Have you ever or are you now serving aboard a CANES ship? [Yes | No | I don't know]
   An answer of "no" or "I don't know" to this question ended the survey.
2. What is the date range in which you served aboard this ship?
3. What is that ship's class? [AS | CG | CVN | DDG | ESB | LCC | LCS | LHA | LHD | LPD | LSD | MCM | PC]
4. What was/is your job title aboard that ship?

The remainder of the questions involve branching, so we illustrate this flow in Figure A.1.

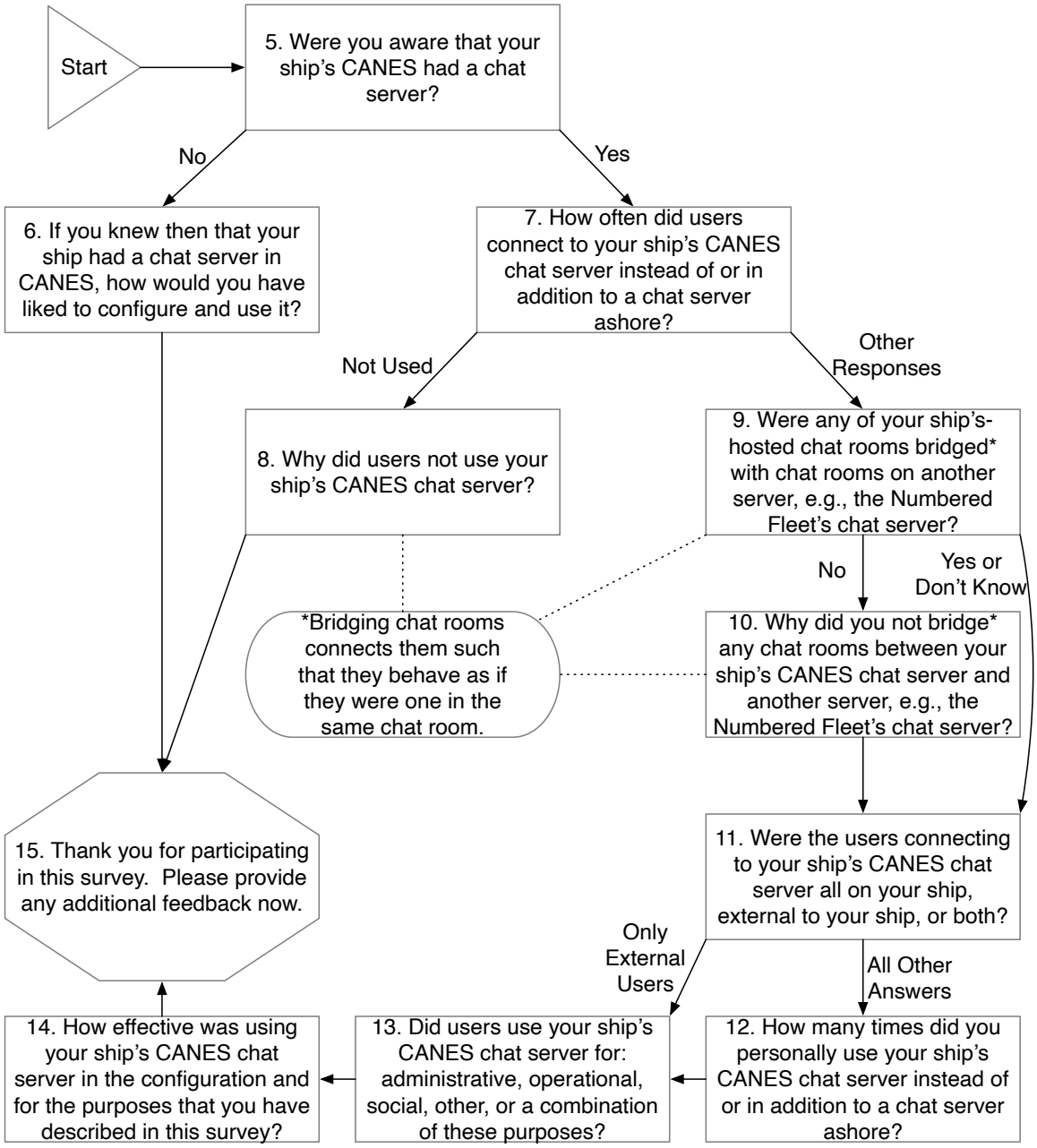


Figure A.1. Chat Server Survey Questions 5–15

The following are the answer options to questions if they are not open-ended or specified in Figure A.1:

7. [We did not use my ship's chat server | Just a few times | Occasionally | Routinely | I don't know]
8. [Select all that apply: I did not know how to connect to it | I did not see a benefit of using it | I did not know that chat rooms could be bridged* with chat rooms hosted on another server, e.g., the Numbered Fleet's chat server | I did not have time to provide support and training to users for how to use it | I don't know | Other (please specify)]
10. [Select all that apply: We did not know this capability existed | We did not have a need to bridge* our locally hosted chat rooms with another server | We could not figure out how to bridge* our chat rooms | We did not take the time to figure out how to bridge* our chat rooms | Other (please specify)]
11. [The users were all on my ship | The users were all external to my ship | Users connected from on my ship and from other ships | I don't know]
12. [I did not use my ship's chat server | Just a few times | Occasionally | Routinely]
13. [Select all that apply: Administrative | Operational | Social | I don't know | Other (please specify)]
14. [Ineffective | Marginally Effective | Moderately Effective | Effective | Highly Effective]

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B:
Fleet Communications Shifts Data Collection Survey

Introduction: You are invited to participate in a research study optimizing fleet Internet Protocol (IP) service communications shifts. The purpose of this research is to find and help the Navy overcome sociotechnical barriers to conducting efficient shifts during changes of operational control (CHOPs) between areas of responsibility (AORs) routinely. The data collection consists of the questions below, estimated to take about 20–30 minutes to complete.

Navy Survey Office Authorization
Control Number NSPM23.23, Expiration date 6/15/2025

Key Information

1. **Participation is voluntary.** Refusal to participate will involve no penalty or loss of benefits to which you would otherwise be entitled, and you may discontinue participation at any time without penalty or loss of benefits to which you otherwise would be entitled.
2. **Risks:** There are no reasonably foreseeable risks or discomforts to prospective subjects.
3. **Benefits:** This data collection supports research into optimizing afloat tactical network communications shifts between areas of responsibility, both from technical and sociotechnical/organizational perspectives. You will not directly benefit from participating in this study.
4. **Study Details:** Up to 17 individuals and ships will participate in this study. The alternative to participating in this study is to not participate.
5. **Who can Participate?** Participation in this study requires your ship to have Automated Digital Network System (ADNS) Inc. III and at least two ADNS-connected satellite communications (SATCOM) missions (not counting extremely high frequency (EHF) Time-division multiple access (TDMA) Interface Processor (TIP)) to shift when conducting a CHOP to another AOR. The intended person to complete this data collection form is the ship's Communications Officer or the Communications Officer's direct

superior because this person likely has the technical data and authority to answer the questions.

6. **Financial:** There is no cost to participate in this research study. You will not be compensated for your participation.
7. **Study Procedure:** The questions below are designed to accompany your ship's existing communications shift procedures for ADNS-connected SATCOM (not including EHF TIP) and the network enclaves directly connected to ADNS. We want to document your pre- and post-shift configurations as well as the order and timing in which you disconnect and (re)connect each SATCOM mission and network enclave. Finally, and just as important as the technical data, we want to know why you decided to disconnect and reconnect in that order as well as information about your experiences with the shifts. These explanations could be technical in nature, based on direction from higher authority, or some combination of both. Please add or remove rows in the tables below as necessary. After the shift, we request that you send this completed survey form to the researcher (specific directions below). If at any point you choose not to participate in this research, simply don't send this survey form.

Confidentiality & Privacy Act: Any information that is obtained during this study will be kept confidential to the full extent permitted by law. All efforts, within reason, will be made to keep your personal information in your research record confidential but total confidentiality cannot be guaranteed. This survey collects only your military rank and job title. However, it is possible that the researcher may be required to divulge information obtained in the course of this research to the subject's chain of command or other legal body but would be limited to the information provided by you to the researcher. If the researcher quotes any portion of your submitted survey form, you will be identified broadly by position, for example, "The Communications Officer (COMMO) on ship #7 said, '. . . ' "

Notes on Classification

- Per the Wideband SATCOM security classification guide (SCG) dated 19 June 2015 (Defense Technical Information Center (DTIC) publication[1] AD1001059), the answers to questions #1 and #2 will be unclassified for most ships. **If you find that**

[1]Department of Defense (DoD) personnel can create an account to access publications online at https://www.dtic.mil/.

your satellite access authorization (SAA) identifies any of these data fields as classified, please do not complete this survey.

- The technical specifications of ADNS are controlled technical information, protected by the DoD's controlled unclassified information (CUI) Program. As such, this form becomes CUI when question #3 is filled in.
- For question #4, please note that the requested date-time groups (DTGs) are your ship's actual disconnect and connect times, which can only exist after the event has occurred and are unclassified. This is in contrast to the times authorized for your ship's SATCOM mission in the SAA, which are often classified and do not necessarily correspond with your actual access times.
- **For your answers to all questions, please keep your explanations at the unclassified level.** The researcher can accept CUI (formerly For Official Use Only (FOUO)) survey responses.

Survey

1. Please list each of the ADNS-connected SATCOM missions (except EHF TIP) before the shift.

Satellite Constellation (e.g., DSCS, WGS, AEHF, CBSP)	STEP Site (e.g., Buckner, Landstuhl, Northwest)	Receive Data Rate in Kbps	Ship's Antenna Group

Note: An antenna group is a linked set of antennas whose azimuth and elevation are controlled by the same controller for tracking a single satellite for the purpose of reducing the size of the block zones caused by the ship's superstructure. Feel free to use any unique identifier for each antenna group. Because ships can have multiple SATCOM missions on one antenna group, all of the SATCOM missions sharing one antenna group necessarily need to be disconnected before the antenna group can be used to connect to a new mission on a different satellite.

2. Please list each of the ADNS-connected SATCOM missions (except EHF TIP) after the shift.

Satellite Constellation (e.g., DSCS, WGS, AEHF, CBSP)	STEP Site (e.g., Buckner, Landstuhl, Northwest)	Receive Data Rate in Kbps	Ship's Antenna Group

Please note that if at least one of the above tables does not have at least two SATCOM missions, your ship is ineligible to participate in this survey.

3. Please list the pre-shift approximate network utilization of each network enclave directly connected to ADNS and, before and after the shift, its servicing network operations center (NOC). Because not all enclaves are available at all NOCs, some enclaves might not shift. In addition, not all ships have the same networks so add, change, or remove rows to suit your configuration.

Network Enclave	Approximate Network Utilization before Shift as Receive Data Rate in Kbps	Servicing NOC (e.g., PRNOC, ECRNOC) before the shift	Servicing NOC (e.g., PRNOC, ECRNOC) after the shift
NIPRNet			
SIPRNet			
JWICS			
CENTRIXS-?			
CENTRIXS-?			
CENTRIXS-?			
CENTRIXS-?			

Note: The utilization metric is available in the ADNS ciphertext (CT) router on your ship. To acquire it, `enable` privileged mode, run the command `show interface summary`, and look at the TXBS (transmit bits per second) for each physical ethernet interface connecting to a plaintext (PT) enclave's tactical local-area network encryptor (TACLANE) in question (e.g., `GigabitEthernet0/0.9` might be for your ship's NIPRNet enclave). To translate from bits/sec to Kbps, divide by 1,000.

4. Please list the order in which you disconnected and connected services from one AOR to the next. Because SATCOM missions usually change significantly, use two different rows in the table. Be aware that some SATCOM shifts occur completely ashore as the NOC patches your connection to a different ADNS Border Router. On the other hand, network enclaves can have just one row for both the disconnection and (re)connection.

SATCOM Mission/Network Enclave (e.g., NIPRNet, CENTRIXS-K, CBSP, WGS #1 8 Mbps)	Disconnect DTG	Connect DTG

5. Please describe why you decided on the order that you listed in the previous question. You could have one or multiple reasons so please 'X' all that apply and provide amplifying information as applicable in place of the "_____" blanks. If you are not the person that made this decision (or recommendation) on what order to complete the shift, please ascertain this information from that person.

X	Reason
__	My ship has (or I have) always shifted our SATCOM and IP network services in this order.
__	The Joint Fleet Telecommunications Watch Officer (JFTOC) (or someone at Naval Computer and Telecommunications Area Master Station (NCTAMS)/Naval Computer and Telecommunications Station (NCTS)) directed my ship to shift services in this order because _____.
__	ADNS or other Naval Information Warfare Center (NIWC) documentation (document #_____) recommends this procedure.
__	My commanding officer (CO) directed us to shift services in this order because _____.
__	My squadron or strike group commander or staff member directed the ship to shift services in this order because _____.
__	The Numbered Fleet Commander (or higher) or staff member directed the ship to shift services in this order because _____.
__	We needed to shift _____ [before/after] _____ because _____ (e.g., scheduled video teleconference (VTC), operational event/task, mission requirement).
__	Any other reason(s): _____.

6. For the previous question, if you were not the person that made this decision (or recommendation) on what order to complete the shift, please enter the job title (e.g., ship's COMMO) and rank (e.g., LT, CIV) of the person that decided on this order?

7. From the perspective of SATCOM and IP network services availability, to what degree are you satisfied with this CHOP: Very Dissatisfied, Dissatisfied, Neither Satisfied nor Dissatisfied, Satisfied, or Very Satisfied?
8. What circumstance(s) led you to select the level of satisfaction that you did in the previous question?
9. To what extent do you see room for improvement in the process of shifting SATCOM and IP network services when conducting a CHOP from one AOR to another: To no extent, To a small extent, To a moderate extent, To a great extent, or To a very great extent?
10. Unless you answered the previous question "to no extent," what ideas do you have for improving the process of shifting SATCOM and IP network services when conducting a CHOP from one AOR to another?
11. As the person completing this survey, what is your rank (e.g., LTJG, CWO2, ITC)?
12. As the person completing this survey, what is your job title (e.g., COMMO)?
13. Please provide any other comments or feedback on the CHOP process for shifting SATCOM and IP network services between AORs that you would like the researchers to consider.

Completion: As a reminder, **all survey data must be unclassified**; the researcher cannot accept any classified information via this survey form. Please submit your completed survey form to the student researcher, CDR Christopher Landis, via this anonymous "file request" on the NPS subdomain of Box, which is accredited to store CUI: [redacted address]. Please do not upload any classified information. If for any reason you determine that it is in your ship's or your best interest not to participate, you can simply refuse to participate by not submitting this completed survey form.

Statement of Agreement: I have read the information provided above. I have been given the opportunity to ask questions and all the questions have been answered to my satisfaction. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research, I do not waive any of my legal rights.

Points of Contact: If you have any questions or comments about the research, or you experience an injury or have questions about any discomforts that you experience while

taking part in this study please contact the Principal Investigator, Dr. Joshua A. Kroll, Faculty Advisor, 831-656-3988, jkroll@nps.edu. Questions about your rights as a research subject or any other concerns may be addressed to the Navy Postgraduate School IRB Chair, Dr. Larry Shattuck, 831-656-2473, lgshattu@nps.edu.

Thank you, again, for your time and support of this research.

Acronyms Used in this Survey:
[All are defined in-line or incorporated in the list on page xiii.]

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C:
Email RFI to USG Federal Agencies

To whom it may concern at the [Agency/Organization],

I am a U.S. Naval Officer working toward a Computer Science PhD at the Naval Postgraduate School (NPS) in Monterey, CA, and am requesting your response to this request for information (RFI) in support of my research.

I am researching how organizational factors affect the formation and implementation of policy in cybersecurity and privacy. As one of my case studies, I am looking into the Department of Homeland Security Binding Operational Directive (BOD) 18-01 (https://www.cisa.gov/binding-operational-directive-18-01), specifically the email security provisions for Sender Policy Framework (SPF) and Domain-based Message Authentication, Reporting, and Conformance (DMARC). An online query of public Domain Name System (DNS) records detected the following:

- Domain	Nonconforming Property
- [List Specific to Recipient]	

I recognize that every organization has unique needs and that a one-size-fits-all approach to cybersecurity is not always effective and am wondering why this domain does not conform with the details of BOD 18-01. For each domain identified, please identify the reason(s) for your organization's approach and unique considerations. For example, these reasons could be technical barriers, policy barriers, manpower issues, staffing/prioritization issues, or other reasons. If you are willing to share this information with me for the benefit of my research, I am interested in knowing why the policy has not been implemented.

I have previously served as a network administrator and so also recognize that competing interests are often at play between cybersecurity objectives, mission requirements, and usability.

If you would like to respond anonymously, you may upload your response as a file to [redacted address]. It is an NPS subdomain of Box and has been accredited to store controlled

unclassified information (CUI), thus meeting the safeguarding requirements of 32 CFR part 2002 (https://www.ecfr.gov/current/title-32/subtitle-B/chapter-XX/part-2002).

Confidentiality: Any information that is obtained during this study will be kept confidential to the full extent permitted by law. As a U.S. Government-operated research university, NPS and its staff (including the researchers) are bound by restrictions on handling sensitive government information. All efforts, within reason, will be made to keep your personal information in your response confidential but total confidentiality cannot be guaranteed. If the researcher quotes any portion of your submission, you will be identified anonymously, for example, "One organization approached this from the perspective of, '. . . ' "

Thank you for your time and support of this research.
[signature block]

APPENDIX D:
Our SPF and DMARC Explanation

Thank you for your response and willingness to support my research.

I have the impression that there may be some confusion regarding the purpose of SPF and DMARC records. SPF and DMARC help you protect your domain against spoofing and impersonation, regardless of whether an email server exists in your domain. When configured properly, others' email servers can verify whether a received email claiming to be from your domain is authorized to have come from your domain and handle it accordingly.

- SPF prevents sender email address domain spoofing by verifying the IP address of the mail server from which it received the email with the specified list of allowed IP addresses in the SPF record. Please see RFC 7208 for more information: https://datatracker.ietf.org/doc/html/rfc7208
- A DMARC record informs servers that have received email from its domain what to do with messages that fail SPF (or DKIM) verification: nothing, quarantine, or reject. It also indicates the recipient for reports on attempts to spoof its domain as an email source. Please see RFC 7489 for more information: https://datatracker.ietf.org/doc/html/rfc7489

Because you indicated that you do not have an email server in your domain, a recommended configuration to disallow email spoofing of your domain include the following SPF and DMARC records:

`"v=spf1 -all"`
. . . as a DNS `TXT` record for [your domain] indicates that "all" servers fail ("-") the test of whether they are authorized to transmit email on behalf of your domain.

`"v=DMARC1; p=reject; pct=100; fo=1;`
`rua=mailto:reports@dmarc.cyber.dhs.gov"`
. . . as a DNS `TXT` record for the _dmarc.[your domain] subdomain communicates to servers receiving email (appearing to be) from your domain that it should "reject" "100" percent

of it that does not pass the SPF (or DKIM) test ("fo=1")[2] and that it should send aggregate reports of these events to the specified email address.

Implementing these policies for your domain would protect against others spoofing it and align your configuration with the SPF and DMARC configurations specified in BOD 18-01.

Thank you, again, for your time and consideration in this matter.

Very respectfully,
Chris Landis

[2] After sending this email, we discovered an error: The failure reporting options (FO) field determines what types of failures should trigger a failure report for sending to the reporting uniform resource identifiers (URIs) for failure data (RUFs). Without an RUF field, the FO field is meaningless [266].

Supplementals

To access the supplemental material(s) listed here, contact the Dudley Knox Library (https://libanswers.nps.edu/) or, for publicly releasable theses and supplementals only, visit the thesis pages in the library's Calhoun database (https://library.nps.edu/nps-theses).

Supplemental #1 of 4: Chat Use Data Collection (CUI-CTI): Sociotechnical Data Collected from the Fleet

Supplemental #2 of 4: Afloat Tactical Network Descriptions (CUI-CTI): Detailed Systems Descriptions Supporting the Research on Communications Shifts and System Developments for the U.S. Navy's Afloat Tactical Networks

Supplemental #3 of 4: Fleet Communications Shifts Data Collection (CUI-CTI): Sociotechnical Data Collected from the Fleet while Deployed

Supplemental #4 of 4: Email Security Query Data (Unclassified): SPF, DMARC, and STARTTLS Data

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

[1] J. Beerman, D. Berent, Z. Falter, and S. Bhunia, "A review of Colonial Pipeline ransomware attack," in *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)*, 2023, pp. 8–15. Available: https://doi.org/10.1109/CCGridW59191.2023.00017

[2] G. A. P. Rodrigues, A. L. M. Serrano, G. F. Vergara, R. d. O. Albuquerque, and G. D. A. Nze, "Impact, compliance, and countermeasures in relation to data breaches in publicly traded U.S. companies," *Future Internet*, vol. 16, no. 6, p. 201, 2024. Available: https://doi.org/10.3390/fi16060201

[3] S. Thielman, "Yahoo hack: 1bn accounts compromised by biggest data breach in history," The Guardian, December 15, 2016. Available: https://www.theguardian.com/technology/2016/dec/14/yahoo-hack-security-of-one-billion-accounts-breached

[4] Federal Bureau of Investigation, "FBI statement on incident involving fake emails," Press Release, November 14, 2021. Available: https://www.fbi.gov/news/press-releases/fbi-statement-on-incident-involving-fake-emails

[5] Federal Bureau of Investigation, "Cyber criminals conduct business email compromise through exploitation of cloud-based email services, costing US businesses more than $2 billion," I-040620-PSA, April 6, 2020. Available: https://www.ic3.gov/Media/Y2020/PSA200406

[6] Federal Trade Commission, "FTC sues Cambridge Analytica, settles with former CEO and app developer," July 24, 2019. Available: https://www.ftc.gov/news-events/news/press-releases/2019/07/ftc-sues-cambridge-analytica-settles-former-ceo-app-developer

[7] Federal Trade Commission, "FTC imposes $5 billion penalty and sweeping new privacy restrictions on Facebook," July 24, 2019. Available: https://www.ftc.gov/news-events/news/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions-facebook

[8] K. Hill, "How Target figured out a teen girl was pregnant before her father did," Forbes, February 16, 2012. Available: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did

[9] R. E. Matland, "Synthesizing the implementation literature: The ambiguity-conflict model of policy implementation," *Journal of Public Administration Research and*

Theory: J-PART, vol. 5, no. 2, pp. 145–174, 1995. Available: https://www.jstor.org/stable/1181674

[10] B. Hudson, D. Hunter, and S. Peckham, "Policy failure and the policy-implementation gap: Can policy support programs help?" *Policy Design and Practice*, vol. 2, no. 1, pp. 1–14, 2019. Available: https://doi.org/10.1080/25741292.2018.1540378

[11] C. I. Barnard, *The Functions of the Executive*. Cambridge, MA: Harvard University Press, 1968. Original work published 1938.

[12] National Institute of Standards and Technology Joint Task Force, *Managing Information Security Risk: Organization, Mission, and Information System View*, Special Publication 800-39, U.S. Department of Commerce. Gaithersburg, MD, March 2011. Available: https://doi.org/10.6028/NIST.SP.800-39

[13] M. C. Bölen, H. Calisir, and Ü. Özen, "Flow theory in the information systems life cycle: The state of the art and future research agenda," *International Journal of Consumer Studies*, vol. 45, no. 4, pp. 546–580, 2021. Available: https://doi.org/10.1111/ijcs.12641

[14] R. S. Dalal, D. J. Howard, R. J. Bennett, C. Posey, S. J. Zaccaro, and B. J. Brummel, "Organizational science and cybersecurity: Abundant opportunities for research at the interface," *Journal of Business and Psychology*, vol. 37, no. 1, pp. 1–29, 2022. Available: https://doi.org/10.1007/s10869-021-09732-9

[15] "Policy," *Dictionary*, 2023. Apple, Inc., Version 2.3.0 (294).

[16] M. J. Maullo and S. B. Calo, "Policy management: An architecture and approach," in *Proceedings of 1993 IEEE 1st International Workshop on Systems Management*, 1993, pp. 13–26. Available: https://doi.org/10.1109/IWSM.1993.315293

[17] J. Feigenbaum and D. J. Weitzner, "On the incommensurability of laws and technical mechanisms: Or, what cryptography can't do," in *Security Protocols XXVI*, 2018, pp. 266–279. Available: https://doi.org/10.1007/978-3-030-03251-7_31

[18] K. Moriarty and S. Farrell, *Deprecating TLS 1.0 and TLS 1.1*, RFC 8996, Internet Engineering Task Force, March 2021. Available: https://datatracker.ietf.org/doc/html/rfc8996

[19] E. C. Duke, *Enhance Email and Web Security*, Binding Operational Directive (BOD) 18-01, U.S. Department of Homeland Security. Washington, DC, October 16, 2017. Available: https://www.cisa.gov/binding-operational-directive-18-01

[20] *DoD Operations Security (OPSEC) Program Manual*, DOD Manual 5205.02, Incorporating Change 2, Under Secretary of Defense for Intelligence, DoD. Washington, DC, 2020. Available: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodm/520502m.PDF

[21] A. van Lamsweerde, *Requirements Engineering: From System Goals to UML Models to Software Specifications*. West Sussex, England: Wiley, 2009.

[22] M. G. Hinchey, J. L. Rash, and C. A. Rouff, "A formal approach to requirements-based programming," in *12th IEEE International Conference and Workshops on the Engineering of Computer-Based Systems (ECBS'05)*, 2005, pp. 339–345. Available: https://doi.org/10.1109/ECBS.2005.7

[23] D. F. Sterne, "On the buzzword "security policy"," in *Proceedings. 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, 1991, pp. 219–230. Available: https://doi.org/10.1109/RISP.1991.130789

[24] M. Backes, M. Durmuth, and G. Karjot, "Unification in privacy policy evaluation – Translating EPAL into Prolog," in *Proceedings. Fifth IEEE International Workshop on Policies for Distributed Systems and Networks, 2004. POLICY 2004.*, 2004, pp. 185–188. Available: https://doi.org/10.1109/POLICY.2004.1309165

[25] G. A. Weaver, "Security-policy analysis with eXtended Unix Tools," Ph.D. dissertation, Dartmouth College, March 1, 2013. Available: https://digitalcommons.dartmouth.edu/dissertations/38/

[26] M. Alohaly, H. Takabi, and E. Blanco, "Automated extraction of attributes from natural language attribute-based access control (ABAC) policies," *Cybersecurity*, vol. 2, no. 1, p. 2, 2019. Available: https://doi.org/10.1186/s42400-018-0019-2

[27] C. Herley, "Unfalsifiability of security claims," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 23, pp. 6415–6420, 2016. Available: https://doi.org/10.2307/26470205

[28] C. Herley, "Justifying security measures — A position paper," in *Computer Security – ESORICS 2017*, 2017, pp. 11–17. Available: https://doi.org/10.1007/978-3-319-66402-6_2

[29] A. Ruighaver, S. Maynard, and S. Chang, "Organisational security culture: Extending the end-user perspective," *Computers & Security*, vol. 26, no. 1, pp. 56–62, 2007. Available: https://doi.org/10.1016/j.cose.2006.10.008

[30] W. J. Triplett, "Addressing human factors in cybersecurity leadership," *Journal of Cybersecurity and Privacy*, vol. 2, no. 3, pp. 573–586, 2022. Available: https://doi.org/10.3390/jcp2030029

[31] P. Carpenter, *Transformational Security Awareness: What Neuroscientists, Storytellers, and Marketers Can Teach Us About Driving Secure Behaviors*. Indianapolis, IN: Wiley, 2019. Available: https://doi.org/10.1002/9781119566380

[32] M. W. Boyce, K. M. Duma, L. J. Hettinger, T. B. Malone, D. P. Wilson, and J. Lockett-Reynolds, "Human performance in cybersecurity: A research agenda," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55, no. 1, pp. 1115–1119, 2011. Available: https://doi.org/10.1177/1071181311551233

[33] A. E. Waldman, *Industry Unbound: The Inside Story of Privacy, Data, and Corporate Power*. Cambridge: Cambridge University Press, 2021. Available: https://doi.org/10.1017/9781108591386

[34] A. Demjaha, S. Parkin, and D. Pym, "You've left me no choices: Security economics to inform behaviour intervention support in organizations," in *Socio-Technical Aspects in Security and Trust*, vol. 11739 (LNCS), T. Groß and T. Tryfonas, Eds. Cham: Springer International Publishing, 2021, pp. 66–86. Available: https://doi.org/10.1007/978-3-030-55958-8_4

[35] C. Herley and P. C. van Oorschot, "SoK: Science, security and the elusive goal of security as a scientific pursuit," in *2017 IEEE Symposium on Security and Privacy (S&P)*, 2017, pp. 99–120. Available: https://doi.org/10.1109/SP.2017.38

[36] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry* (Graduate Studies in Mathematics). Providence, RI: American Mathematical Society, 2001, vol. 33.

[37] K. Forsberg, H. Mooz, and H. Cotterman, *Visualizing Project Management: Models and Frameworks for Mastering Complex Systems*, 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2005.

[38] J. Kouns and D. Minoli, *Information Technology Risk Management in Enterprise Environments: A Review of Industry Practices and a Practical Guide to Risk Management Teams*. Hoboken, NJ: John Wiley & Sons, Inc., 2010.

[39] M. M. Gilday, *Chief of Naval Operations Navigation Plan*, 2022. Available: https://www.dvidshub.net/publication/issues/64582

[40] C. B. Landis, "Afloat resilience of IP tactical communications: Chat systems and sailors," Lightning talk presented at IEEE Conference on Military Communications (MILCOM), Boston, MA, October 31, 2023. DTIC citation AD1215296, restricted distribution. Available: https://milcom2023.milcom.org/program/restricted-technical-program

[41] C. B. Landis and J. A. Kroll, "Acquired and deployed but not adopted: Lost mission effectiveness without resilient chat afloat," in *Proceedings of the Twenty-first Annual Acquisition Research Symposium*, 2024, vol. 3, pp. 199–214. Available: https://dair.nps.edu/handle/123456789/5161

[42] S. Brooks, M. Garcia, N. Lefkovitz, S. Lightman, and E. Nadeau, *An Introduction to Privacy Engineering and Risk Management in Federal Systems*, NISTIR 8062, National Institute of Standards and Technology, U.S. Department of Commerce. Gaithersburg, MD, January 2017. Available: https://doi.org/10.6028/NIST.IR.8062

[43] National Institute of Standards and Technology, *NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management, version 1.0*, U.S. Department of Commerce. Gaithersburg, MD, January 2020. Available: https://doi.org/10.6028/NIST.CSWP.01162020

[44] National Institute of Standards and Technology, "Privacy framework perspectives and success stories," March 22, 2022. Available: https://www.nist.gov/privacy-framework/getting-started-0/perspectives-and-success-stories

[45] C. B. Landis and J. A. Kroll, "Mitigating inference risks with the NIST privacy framework," *Proceedings on Privacy Enhancing Technologies*, vol. 2024, no. 1, pp. 217–231, 2024. Available: https://doi.org/10.56553/popets-2024-0013

[46] National Institute of Standards and Technology Joint Task Force, *Security and Privacy Controls for Information Systems and Organizations*, Special Publication 800-53 Revision 5, U.S. Department of Commerce. Gaithersburg, MD, September 2020. Available: https://doi.org/10.6028/NIST.SP.800-53r5

[47] C. Gonzalez, N. Ben-Asher, A. Oltramari, and C. Lebiere, "Cognition and technology," in *Cyber Defense and Situational Awareness*, A. Kott, C. Wang, and R. F. Erbacher, Eds. Cham: Springer International Publishing, 2014, pp. 93–117. Available: https://doi.org/10.1007/978-3-319-11391-3_6

[48] L. Columbus, "2022 cybersecurity forecasts predict growth, emphasizing resilience," Venture Beat, July 5, 2022. Available: https://venturebeat.com/security/2022-cybersecurity-forecasts-predict-growth-emphasizing-resilience/

[49] A. Wool, "A quantitative study of firewall configuration errors," *Computer*, vol. 37, no. 6, pp. 62–67, June 2004. Available: https://doi.org/10.1109/MC.2004.2

[50] R. Stevens *et al.*, "Compliance cautions: Investigating security issues associated with U.S. digital-security standards," in *Network and Distributed Systems Security (NDSS) Symposium*, February 23, 2020, pp. 1–16. Available: https://doi.org/10.14722/ndss.2020.24003

[51] R. Stevens, J. Dykstra, W. K. Everette, and M. L. Mazurek, "How to hack compliance: Using lessons learned to repeatably audit compliance programs for digital security concerns," in *Learning from Authoritative Security Experiment Results (LASER) 2020*, February 23, 2020, pp. 1–9. Available: https://doi.org/10.14722/laser.2020.23003

[52] L. A. Gunn, "Why is implementation so difficult?" *Management Services in Government*, vol. 33, no. 4, pp. 169–176, 1978. Available: https://archive.org/details/sim_management-in-government_1978-11_33_4/page/168/mode/2up

[53] Assistant Secretary of Defense (Command, Control, Communications, and Intelligence), "Trusted computer system evaluation criteria," Department of Defense, Department of Defense Standard DoD 5200.28-STD, December 26, 1985. Available: http://csrc.nist.gov/publications/history/dod85.pdf

[54] *Federal Information Security Modernization Act of 2014*, Public Law 113-283, United States Congress. Washington, DC, December 18, 2014. Available: https://www.govinfo.gov/content/pkg/PLAW-113publ283/html/PLAW-113publ283.htm

[55] J. W. Meyer and B. Rowan, "Institutionalized organizations: Formal structure as myth and ceremony," *American Journal of Sociology*, vol. 83, no. 2, pp. 340–363, September 1977. Available: https://doi.org/10.1086/226550

[56] S. J. Ali, A. Christin, A. Smart, and R. Katila, "Walking the walk of AI ethics: Organizational challenges and the individualization of risk among ethics entrepreneurs," in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*, 2023, pp. 217–226. Available: https://doi.org/10.1145/3593013.3593990

[57] A. Tilcsik, "From ritual to reality: Demography, ideology, and decoupling in a post-communist government agency," *Academy of Management Journal*, vol. 53, no. 6, pp. 1474–1498, 2010. Available: https://doi.org/10.5465/amj.2010.57318905

[58] K. D. Elsbach and R. I. Sutton, "Acquiring organizational legitimacy through illegitimate actions: A marriage of institutional and impression management theories," *Academy of Management Journal*, vol. 35, no. 4, pp. 699–738, 1992. Available: https://doi.org/10.5465/256313

[59] N. G. Leveson, *Engineering a Safer World*. Cambridge, MA: The MIT Press, 2016. Available: https://doi.org/10.7551/mitpress/8179.001.0001

[60] Y. Wu, W. K. Edwards, and S. Das, "SoK: Social cybersecurity," in *2022 IEEE Symposium on Security and Privacy (S&P)*, 2022, pp. 1863–1879. Available: https://doi.org/10.1109/SP46214.2022.9833757

[61] M. B. Bucks and D. Williams, "Maintaining a competitive advantage: An analysis of the efficiency and effectiveness of the Marine Corps' assessment and authorization process," M.S. thesis, Naval Postgraduate School, Monterey, CA, September 2021. Available: https://hdl.handle.net/10945/68302

[62] D. Ashenden and A. Sasse, "CISOs and organisational culture: Their own worst enemy?" *Computers & Security*, vol. 39, pp. 396–405, 2013. Available: https://doi.org/10.1016/j.cose.2013.09.004

[63] J. Hielscher, U. Menges, S. Parkin, A. Kluge, and M. A. Sasse, ""Employees who don't accept the time security takes are not aware enough": The CISO view of human-centred security," in *32nd USENIX Security Symposium (USENIX Security 23)*, August 2023, pp. 2311–2328. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/hielscher

[64] ASEC's Cybersecurity Working Group, "Description criteria for management's description of an entity's cybersecurity risk management program," American Institute of Certified Public Accountants (AICPA) Assurance Services Executive Committee (ASEC), April 15, 2017. Available: https://www.aicpa.org/resources/download/get-description-criteria-for-a-cybersecurity-risk-management-program

[65] D. Burg, M. Maddison, and R. Watson, "Cybersecurity: How do you rise above the waves of a perfect storm?" Ernst & Young Global Ltd., July 22, 2021. Available: https://www.ey.com/en_gl/cybersecurity/cybersecurity-how-do-you-rise-above-the-waves-of-a-perfect-storm

[66] J. L. Higgs, R. E. Pinsker, T. J. Smith, and G. R. Young, "The relationship between board-level technology committees and reported security breaches," *Journal of Information Systems*, vol. 30, no. 3, pp. 79–98, 2016. Available: https://doi.org/10.2308/isys-51402

[67] S. Garfinkel and H. R. Lipford, *Usable Security: History, Themes, and Challenges* (Synthesis Lectures on Information Security, Privacy, and Trust). Morgan & Claypool Publishers, 2014. Available: https://doi.org/10.2200/S00594ED1V01Y201408SPT011

[68] M. Lennartsson, J. Kävrestad, and M. Nohlberg, "Exploring the meaning of usable security – a literature review," *Information & Computer Security*, vol. 29, no. 4, pp. 647–663, 2021. Available: https://doi.org/10.1108/ICS-10-2020-0167

[69] J. Q. Wilson, *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books, 1989.

[70] A. Coleman *et al.*, "Ambiguity and conflict in policy implementation: The case of the new care models (vanguard) programme in England," *Journal of Social Policy*, vol. 50, no. 2, pp. 285–304, 2021. Available: https://doi.org/10.1017/S0047279420000082

[71] S. Zarb and K. Taylor, "Uneven local implementation of federal policy after disaster: Policy conflict and goal ambiguity," *Review of Policy Research*, vol. 40, no. 1, pp. 63–87, 2023. Available: https://doi.org/10.1111/ropr.12478

[72] R. Wirfs-Brock, "Implementing design responsibilities: Guidelines for constructing usable objects," Wirfs-Brock Associates, White Paper, n.d. Available: https://wirfs-brock.com/PDFs/Implementing%20Design%20Respo.pdf

[73] L. B. Edelman, "Legal ambiguity and symbolic structures: Organizational mediation of civil rights law," *American Journal of Sociology*, vol. 97, no. 6, pp. 1531–1576, May 1992. Available: https://doi.org/10.1086/229939

[74] P. DiMaggio, "The new institutionalisms: Avenues of collaboration," *Journal of Institutional and Theoretical Economics (JITE)*, vol. 154, no. 4, pp. 696–705, 1998. Available: https://www.jstor.org/stable/40752104

[75] E. Skille and C. Stenling, "Inside-out and outside-in: Applying the concept of conventions in the analysis of policy implementation through sport clubs," *International Review for the Sociology of Sport*, vol. 53, no. 7, pp. 837–853, 2018. Available: https://doi.org/10.1177/1012690216685584

[76] R. Ø. Skotnes, "Standardization of cybersecurity for critical infrastructures: The role of sensemaking and translation," in *Standardization and Risk Governance: A Multi-Disciplinary Approach* (Routledge New Security Studies), O. E. Olsen, K. Juhl, P. H. Lindøe, and O. A. Engen, Eds. London: Taylor & Francis, 2020, pp. 166–180. Available: https://doi.org/10.4324/9780429290817-13

[77] C. Perrow, *Normal Accidents: Living with High-Risk Technologies*. Princeton, N.J: Princeton University Press, 1999.

[78] M. C. Elish and T. Hwang, "Praise the machine! Punish the human! The contradictory history of accountability in automated aviation," Intelligence and Autonomy Initiative, May 18, 2015. Available: https://doi.org/10.2139/ssrn.2720477

[79] M. C. Elish, "Moral crumple zones: Cautionary tales in human-robot interaction," *Engaging Science, Technology, and Society*, vol. 5, pp. 40–60, 2019. Available: https://doi.org/10.17351/ests2019.260

[80] B. Green, "The flaws of policies requiring human oversight of government algorithms," *Computer Law & Security Review*, vol. 45, p. 105681, 2022. Available: https://doi.org/10.1016/j.clsr.2022.105681

[81] T. van Steen, "Measuring behavioural cybersecurity: An overview of options," in *Augmented Cognition*, vol. 14019 (LNCS), D. D. Schmorrow and C. M. Fidopiastis, Eds. Cham: Springer Nature Switzerland, 2023, pp. 460–471. Available: https://doi.org/10.1007/978-3-031-35017-7_29

[82] O. Beris, A. Beautement, and M. A. Sasse, "Employee rule breakers, excuse makers and security champions: Mapping the risk perceptions and emotions that drive security behaviors," in *Proceedings of the 2015 New Security Paradigms Workshop (NSPW '15)*, 2015, pp. 73–84. Available: https://doi.org/10.1145/2841113.2841119

[83] R. Anderson, "Why cryptosystems fail," in *Proceedings of the 1st ACM Conference on Computer and Communications Security*, 1993, pp. 215–227. Available: https://doi.org/10.1145/168588.168615

[84] Y. Acar, S. Fahl, and M. L. Mazurek, "You are not your developer, either: A research agenda for usable security and privacy research beyond end users," in *2016 IEEE Cybersecurity Development (SecDev)*, 2016, pp. 3–8. Available: https://doi.org/10.1109/SecDev.2016.013

[85] J. Müller *et al.*, ""Johnny, you are fired!" – Spoofing OpenPGP and S/MIME signatures in emails," in *28th USENIX Security Symposium (USENIX Security 19)*, 2019, pp. 1011–1028. Available: https://www.usenix.org/conference/usenixsecurity19/presentation/muller

[86] L. Bainbridge, "Ironies of automation," *IFAC Proceedings Volumes: IFAC/IFIP/IFORS/IEA Conference on Analysis, Design and Evaluation of Man-Machine Systems, Baden-Baden, FRG, 27-29 September*, vol. 15, no. 6, pp. 129–135, September 1982. Available: https://doi.org/10.1016/S1474-6670(17)62897-0

[87] M. Mazurek, "We are the experts, and we are the problem: The security advice fiasco," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 2022, p. 7. Available: https://doi.org/10.1145/3548606.3559394

[88] J. McAlaney and V. Benson, "Cybersecurity as a social phenomenon," in *Cyber Influence and Cognitive Threats*, V. Benson and J. McAlaney, Eds. Academic Press, 2020, pp. 1–8. Available: https://doi.org/10.1016/B978-0-12-819204-7.00001-4

[89] A. B. Haynes *et al.*, "A surgical safety checklist to reduce morbidity and mortality in a global population," *New England Journal of Medicine*, vol. 360, no. 5, pp. 491–499, January 2009. Available: https://doi.org/10.1056/NEJMsa0810119

[90] D. Maclean, "The NIST Risk Management Framework: Problems and recommendations," *Cyber Security: A Peer-Reviewed Journal*, vol. 1, no. 3, pp. 207–217, 2017. Available: https://www.ingentaconnect.com/content/hsp/jcs/2017/00000001/00000003/art00003

[91] S. Bellovin, "Security by checklist," *IEEE Security & Privacy*, vol. 6, no. 2, p. 88, 2008. Available: https://doi.org/10.1109/MSP.2008.43

[92] R. Stevens, J. Dykstra, W. Knox Everette, and M. L. Mazurek, "It lurks within: A look at the unexpected security implications of compliance programs," *IEEE Security & Privacy*, vol. 18, no. 6, pp. 51–58, 2020. Available: https://doi.org/10.1109/MSEC.2020.3014291

[93] A. Gawande, *The Checklist Manifesto: How to Get Things Right*. New York: Metropolitan Books, 2010.

[94] J. G. Schwitz, "Risk-based cybersecurity policy," *American Intelligence Journal*, vol. 29, no. 1, pp. 115–125, 2011. Available: https://www.jstor.org/stable/26201928

[95] S. S. Silbey, "Taming Prometheus: Talk about safety and culture," *Annual Review of Sociology*, vol. 35, no. 1, pp. 341–369, 2009. Available: https://doi.org/10.1146/annurev.soc.34.040507.134707

[96] D. Branley-Bell, L. Coventry, and E. Sillence, "Promoting cybersecurity culture change in healthcare," in *The 14th PErvasive Technologies Related to Assistive Environments Conference (PETRA '21)*, 2021, pp. 544–549. Available: https://doi.org/10.1145/3453892.3461622

[97] M. Alshaikh and B. Adamson, "From awareness to influence: Toward a model for improving employees' security behaviour," *Personal and Ubiquitous Computing*, vol. 25, no. 5, pp. 829–841, October 2021. Available: https://doi.org/10.1007/s00779-021-01551-2

[98] J. M. Haney and W. G. Lutters, "Motivating cybersecurity advocates: Implications for recruitment and retention," in *Proceedings of the 2019 on Computers and People Research Conference*, 2019, pp. 109–117. Available: https://doi.org/10.1145/3322385.3322388

[99] R. Connolly, "Why computing belongs within the social sciences," *Communications of the ACM*, vol. 63, no. 8, pp. 54–59, July 2020. Available: https://doi.org/10.1145/3383444

[100] A. Beautement, M. A. Sasse, and M. Wonham, "The compliance budget: Managing security behaviour in organisations," in *Proceedings of the 2008 New Security*

Paradigms Workshop (NSPW '08), 2008, pp. 47–58. Available: https://doi.org/10.1145/1595676.1595684

[101] B. Latour, "Where are the missing masses? The sociology of a few mundane artifacts," in *Shaping Technology/Building Society: Studies in Sociotechnical Change*, W. E. Bijker and J. Law, Eds. Cambridge, MA: MIT Press, 1992, pp. 225–258. Available: https://ieeexplore.ieee.org/document/9592179

[102] J. R. Reidenberg, "Lex informatica: The formulation of information policy rules through technology," *Texas Law Review*, vol. 76, no. 3, pp. 553–593, 1998. Available: https://www.proquest.com/docview/203689546

[103] L. Lessig, *Code: And Other Laws of Cyberspace*. New York: Basic Books, 1999.

[104] B. J. Chen and J. Metcalf, "Explainer: A sociotechnical approach to AI policy," Data & Society, Policy Brief, May 2024. Available: https://datasociety.net/library/a-sociotechnical-approach-to-ai-policy/

[105] A. Z. Jacobs, J. A. Kroll, R. S. Geiger, and A. Smart, "Unsafe at any AUC: Unlearned lessons from sociotechnical disasters for responsible AI," 2024, unpublished.

[106] D. Vaughan, "Rational choice, situated action, and the social control of organizations," *Law & Society Review*, vol. 32, no. 1, pp. 23–61, 1998. Available: https://doi.org/10.2307/827748

[107] D. Vaughan, "Organizational rituals of risk and error," in *Organizational Encounters with Risk*, B. Hutter and M. Power, Eds. New York: Cambridge University Press, 2005, pp. 33–66.

[108] D. A. Norman, *The Design of Everyday Things*, revised and expanded ed. New York: Basic Books, 2013.

[109] S. L. Star, "The ethnography of infrastructure," *American Behavioral Scientist*, vol. 43, no. 3, pp. 377–391, 1999. Available: https://doi.org/10.1177/00027649921955326

[110] O. Martin and A. Marcley, "Mako chat support," NIWC Pacific, San Diego, CA, Notice of Intent N66001_SNOTE_0007EFDE, July 9, 2013. Available: https://sam.gov/opp/08d1ade640cd6fba0a02462113dbc880/view

[111] B. A. Eovito, "An assessment of joint chat requirements from current usage patterns," M.S. thesis, Naval Postgraduate School, Monterey, CA, June 2006. Available: https://hdl.handle.net/10945/2753

[112] D. A. DeVos, "XML tactical chat (XTC) the way ahead for navy chat," M.S. thesis, Naval Postgraduate School, Monterey, CA, September 2007. Available: https://hdl.handle.net/10945/3312

[113] N. J. Heacox, R. A. Moore, J. G. Morrison, and R. Yturralde, "Real-time online communications: 'Chat' use in navy operations," Pacific Science & Engineering Group, Inc., San Diego, CA, Report, June 2004. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=be4d48d345e1cc87593cfb96c4258987085ad03d

[114] D. Barrett, "Carrier strike group twelve sponsors fleet/joint/coalition testing of open standards chat tool," *CHIPS*, January 2006. Available: https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=3148

[115] J. Catanzaro, "User-centered design requirements for maritime chat clients in network-centric warfare," Pacific Science & Engineering Group, Inc., San Diego, CA, Report, 2006. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4c60f78f80a247e49bb044b0a802efbb78ed081e

[116] C. B. Landis, "A new era of afloat IP services," *CHIPS*, October 2016. Available: https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=8349

[117] M. M. Morris, L. E. Banchs, and K. R. Stout, "Delivering assured command and control through wide area network optimization," M.S. thesis, Naval Postgraduate School, Monterey, CA, June 2024. Restricted distribution.

[118] N. M. Simila and W. E. Tooman, "Assured command and control thin line automation via quality-of-service refinement," M.S. thesis, Naval Postgraduate School, Monterey, CA, June 2022. Restricted distribution.

[119] D. P. Bartlett, "Satellite mesh architectures in the maritime domain," M.S. thesis, Naval Postgraduate School, Monterey, CA, March 2014. Restricted distribution.

[120] V. Machi, "Navy beefing up at-sea enterprise network," *National Defense*, vol. 102, no. 772, pp. 30–31, 2018. Available: https://www.jstor.org/stable/27022127

[121] D. A. Broniatowski, "Building the tower without climbing it: Progress in engineering systems," *Systems Engineering*, vol. 21, no. 3, pp. 259–281, 2018. Available: https://doi.org/10.1002/sys.21426

[122] D. L. Boslaugh, *When Computers Went to Sea: The Digitization of the United States Navy*. Los Alamitos, CA: IEEE Computer Society, 1999.

[123] B. Strauch, "Ironies of automation: Still unresolved after all these years," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 5, pp. 419–433, 2018. Available: https://doi.org/10.1109/THMS.2017.2732506

[124] J. A. Sullivan, "Management of autonomous systems in the Navy's Automated Digital Network System (ADNS)," M.S. thesis, Naval Postgraduate School, Monterey, CA, September 1997. Available: https://hdl.handle.net/10945/8645

[125] R. Fernandez and D. Rodrik, "Resistance to reform: Status quo bias in the presence of individual-specific uncertainty," *The American Economic Review*, vol. 81, no. 5, pp. 1146–1155, 1991. Available: https://www.jstor.org/stable/2006910

[126] B. Hofman, G. de Vries, and G. van de Kaa, "Keeping things as they are: How status quo biases and traditions along with a lack of information transparency in the building industry slow down the adoption of innovative sustainable technologies," *Sustainability*, vol. 14, no. 13, 2022. Available: https://doi.org/10.3390/su14138188

[127] G. Wipfler. *2023 Spring Quarter Graduation - June 16, 2023 (full)*. (June 16, 2023). 29:20–43:23. Available: https://www.youtube.com/watch?v=Xuw8zqsPRTQ

[128] W. J. Orlikowski, J. Yates, K. Okamura, and M. Fujimoto, "Shaping electronic communication: The metastructuring of technology in the context of use," *Organization Science*, vol. 6, no. 4, pp. 423–444, 1995. Available: https://doi.org/10.1287/orsc.6.4.423

[129] G. DeSanctis and M. S. Poole, "Capturing the complexity in advanced technology use: Adaptive structuration theory," *Organization Science*, vol. 5, no. 2, pp. 121–147, 1994. Available: https://doi.org/10.1287/orsc.5.2.121

[130] S. R. Barley, "Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments," *Administrative Science Quarterly*, vol. 31, no. 1, pp. 78–108, 1986. Available: https://doi.org/10.2307/2392767

[131] N. Sikinger, "IP community health brief," My Navy HR, September 30, 2022. Restricted distribution.

[132] "Navy Information Professionals," milSuite, n.d. Access is restricted to Department of Defense Common Access Card holders. Available: https://www.milsuite.mil/book/groups/navy-information-professionals

[133] "Flank speed – Navy's transition to improved Microsoft 365 cloud collaboration," U.S. Navy Press Release, May 11, 2021. Available: https://www.navy.mil/Press-Office/Press-Releases/display-pressreleases/Article/2602863/flank-speed-navys-transition-to-improved-microsoft-365-cloud-collaboration/

[134] M. D. C. Tongco, "Purposive sampling as a tool for informant selection," *Ethnobotany Research & Applications*, vol. 5, pp. 147–158, 2007. Available: https://hdl.handle.net/10125/227

[135] *Mako 2.0: User Guide*, Version 1.46, PEO C4I, April 26, 2021. Restricted distribution.

[136] *Mako 2.0: Administrator Guide*, Version 1.53, PEO C4I, April 26, 2021. Restricted distribution.

[137] D. H. Anunciado, personal communication, October 3, 2023.

[138] Government Accountability Office, "Immigration service: INS' technology selection process is weak, informal, and inconsistently applied," Washington, DC, Report GAO/PMED-88-16, April 22, 1988. Available: https://www.gao.gov/products/pemd-88-16

[139] Department of the Navy, "Catalog of navy training courses (CANTRAC)," n.d. Accessed on June 30, 2023. Access is restricted to Department of Defense Common Access Card holders. Available: https://app.prod.cetars.training.navy.mil/cantrac/vol2.html

[140] C. R. B. de Souza, C. S. Pinhanez, and V. F. Cavalcante, "Information needs of system administrators in information technology service factories," in *Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology (CHIMIT '11)*, 2011, pp. 1–10. Available: https://doi.org/10.1145/2076444.2076447

[141] W. D. Rowe, "Understanding uncertainty," *Risk Analysis*, vol. 14, no. 5, pp. 743–750, 1994. Available: https://doi.org/10.1111/j.1539-6924.1994.tb00284.x

[142] E. Johnson, "USS Enterprise SHF cross-connect configuration increases allocated bandwidth by 100 percent," *CHIPS*, April 2011. Available: https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=2298

[143] S. Blanchette, Jr., C. Albert, and S. Garcia-Miller, "Beyond technology readiness levels for software: U.S. Army workshop report," Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, Technical Report CMU/SEI-2010-TR-044, December 2010. Available: https://apps.dtic.mil/sti/citations/tr/ADA535517

[144] *Operation of the Software Acquisition Pathway*, DoD Instruction 5000.87, Department of Defense. Washington, DC, October 2, 2020. Available: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500080p.PDF

[145] J. Bone, "Cognitive risk framework for cybersecurity: Redesigning risk management and internal controls design," in *Cognitive Hack: The New Battleground in Cybersecurity. . . the Human Mind*. New York: Auerbach Publications, 2017, pp. 133–163. Available: https://doi.org/10.1201/9781315368412-6

[146] J. Springs and M. Weaver, "Watch turnover support for future surface combatants," *Naval Engineers Journal*, vol. 115, no. 2, pp. 75–88, 2003. Available: https://doi.org/10.1111/j.1559-3584.2003.tb00206.x

[147] M. Fimin, "Breaking bad: Avoiding the 10 worst IT admin habits," *Network Security*, vol. 2016, no. 1, pp. 8–11, 2016. Available: https://doi.org/10.1016/S1353-4858(16)30007-1

[148] L. Vinsel and A. L. Russell, *The Innovation Delusion: How Our Obsession with the New Has Disrupted the Work That Matters Most*. New York: Currency, 2020.

[149] S. Ruggles and D. L. Magnuson, ""It's none of their damn business": Privacy and disclosure control in the U.S. census, 1790–2020," *Population and Development Review*, vol. 49, no. 3, pp. 651–679, September 2023. Available: https://doi.org/10.1111/padr.12580

[150] S. L. Garfinkel, J. M. Abowd, and S. Powazek, "Issues encountered deploying differential privacy," in *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, 2018, pp. 133–137. Available: https://doi.org/10.1145/3267323.3268949

[151] D. E. Denning and J. Schlorer, "Inference controls for statistical databases," *Computer*, vol. 15, no. 7, pp. 69–82, July 1983. Available: https://doi.org/10.1109/MC.1983.1654444

[152] N. R. Adam and J. C. Worthmann, "Security-control methods for statistical databases: A comparative study," *ACM Computing Surveys (CSUR)*, vol. 21, no. 4, pp. 515–556, December 4, 1989. Available: https://doi.org/10.1145/76894.76895

[153] A. Cohen and K. Nissim, "Towards formalizing the GDPR's notion of singling out," *Proceedings of the National Academy of Sciences*, vol. 117, no. 15, pp. 8344–8352, 2020. Available: https://doi.org/10.1073/pnas.1914598117

[154] M. Elliot *et al.*, "Functional anonymisation: Personal data and the data environment," *Computer Law & Security Review*, vol. 34, no. 2, pp. 204–221, 2018. Available: https://doi.org/10.1016/j.clsr.2018.02.001

[155] J. Powar and A. R. Beresford, "SoK: Managing risks of linkage attacks on data privacy," *Proceedings on Privacy Enhancing Technologies*, vol. 2023, no. 2, pp. 97–116, 2023. Available: https://doi.org/10.56553/popets-2023-0043

[156] K. A. Bamberger and D. K. Mulligan, *Privacy on the Ground: Driving Corporate Behavior in the United States and Europe*. Cambridge, MA: MIT Press, 2015.

[157] D. A. Shehab and M. J. Alhaddad, "Comprehensive survey of multimedia steganalysis: Techniques, evaluations, and trends in future research," *Symmetry*, vol. 14, no. 1, p. 117, 2022. Available: https://doi.org/10.3390/sym14010117

[158] P. M. Schwartz and D. J. Solove, "The PII problem: Privacy and a new concept of personally identifiable information," *New York University Law Review (1950)*, vol. 86, no. 6, pp. 1814–1894, 2011. Available: https://www.law.berkeley.edu/files/bclt_Schwartz-Solove_NYU_Final_Print.pdf

[159] P. Ohm, "Broken promises of privacy: Responding to the surprising failure of anonymization," *UCLA Law Review*, vol. 57, no. 6, pp. 1701–1777, 2010. Available: https://www.uclalawreview.org/pdf/57-6-3.pdf

[160] A. Cohen, "Attacks on deidentification's defenses," in *31st USENIX Security Symposium (USENIX Security 22)*, August 2022, pp. 1469–1486. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/cohen

[161] K. Nissim *et al.*, "Bridging the gap between computer science and legal approaches to privacy," *Harvard Journal of Law & Technology*, vol. 31, no. 2, pp. 687–780, 2018. Available: https://jolt.law.harvard.edu/assets/articlePDFs/v31/02.-Article-Wood-7.21.pdf

[162] C. Dwork and A. Roth, *The Algorithmic Foundations of Differential Privacy* (Foundations and Trends® in Theoretical Computer Science). Boston, MA: Now Publishers, Inc., 2014, vol. 9, no. 3–4. Available: https://doi.org/10.1561/0400000042

[163] I. Goldberg, "Privacy-enhancing technologies for the internet III: Ten years later," in *Digital Privacy: Theory, Technologies, and Practices*, A. Acquisti, S. Gritzalis, C. Lambrinoudakis, and S. di Vimercati, Eds. New York: Auerbach Publications, 2007, ch. 1, pp. 3–18. Available: https://doi.org/10.1201/9781420052183-7

[164] I. Wagner and D. Eckhoff, "Technical privacy metrics: A systematic survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 3, pp. 1–38, June 2018. Available: https://doi.org/10.1145/3168389

[165] National Institute of Standards and Technology, *Cybersecurity Framework/Privacy Framework to NIST Special Publication 800-53, Revision 5 Mapping*, U.S. Department of Commerce. Gaithersburg, MD, December 2020. Available: https://www.nist.gov/privacy-framework/nist-privacy-framework-and-cybersecurity-framework-nist-special-publication-800-53

[166] J. A. Kroll, N. Kohli, and P. Laskowski, "Privacy and policy in polystores: A data management research agenda," in *Heterogeneous Data Management, Polystores, and Analytics for Healthcare*, vol. 11721 (LNCS), V. Gadepally *et al.*, Eds. Cham: Springer, 2019, pp. 68–81. Available: https://doi.org/10.1007/978-3-030-33752-0_5

[167] S. M. Bellovin, R. M. Hutchins, T. Jebara, and S. Zimmeck, "When enough is enough: Location tracking, mosaic theory, and machine learning," *New York University Journal of Law & Liberty*, vol. 8, pp. 555–628, 2014. Available: https://digitalcommons.law.umaryland.edu/fac_pubs/1375/

[168] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *2008 IEEE Symposium on Security and Privacy (S&P)*, 2008, pp. 111–125. Available: https://doi.org/10.1109/SP.2008.33

[169] M. S. Ferdous, S. Chowdhury, and J. M. Jose, "Privacy threat model in lifelogging," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, 2016, pp. 576–581. Available: https://doi.org/10.1145/2968219.2968324

[170] R. Reagan, *National Operations Security Program*, National Security Decision Directive Number 298, The White House, January 22, 1988. Available: https://catalog.archives.gov/id/6879871

[171] A. Solow-Niederman, "Information privacy and the inference economy," *Northwestern University Law Review*, vol. 117, no. 2, pp. 1–68, 2022. Available: https://doi.org/10.2139/ssrn.3921003

[172] C. P. Hoffmann, C. Lutz, and G. Ranzini, "Privacy cynicism: A new approach to the privacy paradox," *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 10, no. 4, p. article 7, December 6, 2016. Available: https://doi.org/10.2139/ssrn.3319830

[173] L. Sweeney, "*k*-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002. Available: https://doi.org/10.1142/S0218488502001648

[174] D. K. Mulligan, C. Koopman, and N. Doty, "Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2083, p. 20160118, 2016. Available: https://doi.org/10.1098/rsta.2016.0118

[175] C. Dwork and D. K. Mulligan, "It's not privacy, and it's not fair," *Stanford Law Review Online*, vol. 66, no. 35, pp. 35–40, September 2013. Available: https://stanfordlawreview.org/wp-content/uploads/sites/3/2016/08/DworkMullliganSLR.pdf

[176] *Family Educational Rights and Privacy Act of 1974*, 20 U.S.C. Section 1232g, United States Congress. Washington, DC, January 14, 2013. Available: https://www.law.cornell.edu/uscode/text/20/1232g

[177] A. Narayanan and E. W. Felten, "No silver bullet: De-identification still doesn't work," July 9, 2014, Unpublished. Available: https://www.cs.princeton.edu/~arvindn/publications/no-silver-bullet-de-identification.pdf

[178] A. Cavoukian and D. Castro, *Big Data and Innovation, Setting the Record Straight: De-identification* Does *Work*. Ontario, Canada: Office of the Information and Privacy Commissioner, June 16, 2014. Available: https://www2.itif.org/2014-big-data-deidentification.pdf

[179] A. Narayanan, J. Huey, and E. W. Felten, "A precautionary approach to big data privacy," in *Data Protection on the Move: Current Developments in ICT and Privacy/Data Protection*, S. Gutwirth, R. Leenes, and P. De Hert, Eds. Springer, 2016, pp. 357–385. Available: https://doi.org/10.1007/978-94-017-7376-8_13

[180] E. Young, "Educational privacy in the online classroom: FERPA, MOOCs, and the big data conundrum," *Harvard Journal of Law & Technology*, vol. 28, no. 2, pp. 549–592, 2015. Available: https://jolt.law.harvard.edu/articles/pdf/v28/28HarvJLTech549.pdf

[181] M. Douriez, H. Doraiswamy, J. Freire, and C. T. Silva, "Anonymizing NYC taxi data: Does it matter?" in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2016, pp. 140–148. Available: https://doi.org/10.1109/DSAA.2016.21

[182] S. Harris, *CISSP All-In-One Exam Guide*, 6th ed. Berkeley: Osborne, 2012.

[183] A. Tockar, "Riding with the stars: Passenger privacy in the NYC taxicab dataset," Wordpress, September 2014. Available: https://agkn.wordpress.com/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/

[184] G. A. Fowler, "What does your car know about you? We hacked a Chevy to find out," The Washington Post, December 17, 2019. Available: https://www.washingtonpost.com/technology/2019/12/17/what-does-your-car-know-about-you-we-hacked-chevy-find-out/

[185] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the crowd: The privacy bounds of human mobility," *Scientific Reports*, vol. 3, no. 1, pp. 1376–1376, 2013. Available: https://doi.org/10.1038/srep01376

[186] J. Caltrider, M. Rykov, and Z. MacDonald, "It's official: Cars are the worst product category we have ever reviewed for privacy," Mozilla, September 6, 2023. Available: https://foundation.mozilla.org/en/privacynotincluded/articles/its-official-cars-are-the-worst-product-category-we-have-ever-reviewed-for-privacy/

[187] A. Hern, "Strava suggests military users 'opt out' of heatmap as row deepens," The Guardian, January 29, 2018. Available: https://www.theguardian.com/technology/2018/jan/29/strava-secret-army-base-locations-heatmap-public-users-military-ban

[188] A. Hern, "Fitness tracking app Strava gives away location of secret US Army bases," The Guardian, January 28, 2018. Available: https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

[189] S. Schafer, "With capital in panic, pizza deliveries soar," *Washington Post*, p. D1, December 19, 1998. Available: https://www.washingtonpost.com/wp-srv/politics/special/clinton/stories/pizza121998.htm

[190] M. Barbaro and T. Zeller, "A face is exposed for AOL searcher no. 4417749," The New York Times, August 9, 2006. Available: https://www.nytimes.com/2006/08/09/technology/09aol.html

[191] S. Schoen, "What information is "personally identifiable"?" Electronic Frontier Foundation (EFF), September 11, 2009. Available: https://www.eff.org/deeplinks/2009/09/what-information-personally-identifiable

[192] M. Thalen, "Hacker reveals smart meters are spilling secrets about the Texas snowstorm," Daily Dot, June 24, 2021. Available: https://www.dailydot.com/debug/hacker-smart-meter-texas-snowstorm/

[193] I. Lapowsky, "How Cambridge Analytica sparked the great privacy awakening," Wired, March 2019. Available: https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

[194] *Health Insurance Portability and Accountability Act of 1996*, Public Law 104-191, United States Congress. Washington, DC, August 21, 1996. Available: https://www.govinfo.gov/link/plaw/104/public/191?link-type=pdf&.pdf

[195] U.S. Department of Health and Human Services, "Health information privacy beyond HIPAA: A 2018 environmental scan of major trends and challenges,"

National Committee on Vital and Health Statistics (NCVHS) and its Privacy, Security, and Confidentiality Subcommittee, Rep., December 13, 2017. Available: https://www.ncvhs.hhs.gov/wp-content/uploads/2018/02/NCVHS-Beyond-HIPAA_Report-Final-02-08-18.pdf

[196] *Children's Online Privacy Protection Act of 1998*, 15 U.S.C. Ch. 91, United States Congress. Washington, DC, October 21, 1998. Available: https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter91&edition=prelim

[197] S. Kaufman, "The invisible, yet omnipresent ear: The insufficiencies of the children's online privacy protection act," *NYU Annual Survey of American Law*, vol. 78, pp. 101–136, 2022. Available: https://annualsurveyofamericanlaw.org/wp-content/uploads/2022/11/Kaufman-78.1.pdf

[198] A. E. Holder, "What we don't know they know: What to do about inferences in European and California data protection law," *Berkeley Technology Law Journal*, vol. 35, no. 4, pp. 1331–1364, 2020. Available: https://doi.org/10.15779/Z38MP4VP1V

[199] D. Clifford, M. Richardson, and N. Witzleb, "Artificial intelligence and sensitive inferences: New challenges for data protection laws," in *Regulatory Insights on Artificial Intelligence: Research for Policy*, M. Findlay, J. Ford, J. Seah, and D. Thampapillai, Eds. Cheltenham, UK: Edward Elgar Publishing, 2022, pp. 19–45. Available: https://doi.org/10.4337/9781800880788.00008

[200] F. Feng, X. Wang, and T. Chen, "Analysis of the attributes of rights to inferred information and China's choice of legal regulation," *Computer Law & Security Review*, vol. 41, p. 105565, 2021. Available: https://doi.org/10.1016/j.clsr.2021.105565

[201] D. J. Solove and D. K. Citron, "Risk and anxiety: A theory of data-breach harms," *Texas Law Review*, vol. 96, pp. 737–786, 2018. Available: https://doi.org/10.2139/ssrn.2885638

[202] S. Wachter and B. Mittelstadt, "A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI," *Columbia Business Law Review*, vol. 2019, no. 2, pp. 494–620, 2019. Available: https://doi.org/10.7916/cblr.v2019i2.3424

[203] C. Duhigg, "How companies learn your secrets," The New York Times Magazine, February 16, 2012. Available: https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=1&hp

[204] J. Vertesi, "My experiment opting out of big data made me look like a criminal," Time, May 1, 2014. Available: https://time.com/83200/privacy-internet-big-data-opt-out/

[205] L. Edwards and M. Veale, "Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for," *Duke Law and Technology Review*, vol. 16, no. 1, pp. 18–84, 2017. Available: https://discovery.ucl.ac.uk/id/eprint/1574817/1/Veale_slavetothealgorithm_published.pdf

[206] *California Consumer Privacy Act of 2018*, State of California, 2018. Available: https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

[207] R. Bonta and S. D. Lee. (2022, March 10). Office of the Attorney General, State of California, Opinion. Office of the Attorney General, State of California. No. 20-303. Available: https://oag.ca.gov/system/files/opinions/pdfs/20-303.pdf

[208] S. Nakajima, "Digital advertising: The Google/Facebook duopoly continues," in *Digiworld Yearbook*. Montpellier: Institut de l'Audiovisuel et de Telecommunications en Europe (IDATE), 2019, pp. 80–81. Available: https://www.proquest.com/docview/2445995003/

[209] P. McGee, "Google advertising boom lifts profits to record," *Financial Times*, April 27, 2021. Available: https://www.proquest.com/docview/2532469468

[210] M. Brettel, J.-C. Reich, J. M. Gavilanes, and T. C. Flatten, "What drives advertising success on Facebook? An advertising-effectiveness model measuring the effects on sales of "likes" and other social-network stimuli," *Journal of Advertising Research*, vol. 55, no. 2, pp. 162–175, 2015. Available: https://doi.org/10.2501/JAR-55-2-162-175

[211] I. Sivan-Sevilla, W. Chu, X. Liang, and H. Nissenbaum, "Unaccounted privacy violation," in *FTC PrivacyCon 2020*, 2020, pp. 1–25. Available: https://www.ftc.gov/system/files/documents/public_events/1548288/privacycon-2020-ido_sivan-sevilla.pdf

[212] J. McLeod, "Double-double tracking: How Tim Hortons knows where you sleep, work and vacation," Financial Post, June 12, 2020. Available: https://financialpost.com/technology/tim-hortons-app-tracking-customers-intimate-data

[213] C. Oh, C. Kanich, D. McCoy, and P. Pearce, "Cart-ology: Intercepting targeted advertising via ad network identity entanglement," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 2022, pp. 2401–2414. Available: https://doi.org/10.1145/3548606.3560641

[214] A. Alcaide, E. Palomar, I. Barroso-Perez, and A. E. Abdallah, "Privacy-enhancing cryptography-based materials," in *Proceedings of the International Conference on Security and Cryptography*, 2011, pp. 379–382. Available: https://ieeexplore.ieee.org/document/6732417

[215] L. T. A. N. Brandão and R. Peralta, "Privacy-enhancing cryptography to complement differential privacy," PEC Project, Cryptographic Technology Group, NIST, Post #11, November 3, 2021. Available: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=932789

[216] D. Desfontaines and B. Pejó, "SoK: Differential privacies," in *Proceedings on Privacy Enhancing Technologies*, no. 2, 2020, pp. 288–313. Available: https://doi.org/10.2478/popets-2020-0028

[217] A. A. Abdu, L. M. Chambers, D. K. Mulligan, and A. Z. Jacobs, "Algorithmic transparency and participation through the handoff lens: Lessons learned from the U.S. Census Bureau's adoption of differential privacy," in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, 2024, pp. 1150–1162. Available: https://doi.org/10.1145/3630106.3658962

[218] X. Yi, R. Paulet, and E. Bertino, "Homomorphic encryption," in *Homomorphic Encryption and Applications*. Cham: Springer, 2014, pp. 27–46. Available: https://doi.org/10.1007/978-3-319-12229-8_2

[219] C. Marcolla, V. Sucasas, M. Manzano, R. Bassoli, F. H. P. Fitzek, and N. Aaraj, "Survey on fully homomorphic encryption, theory, and applications," *Proceedings of the IEEE*, vol. 110, no. 10, pp. 1572–1609, 2022. Available: https://doi.org/10.1109/JPROC.2022.3205665

[220] E. Shi, W. Aqeel, B. Chandrasekaran, and B. Maggs, "Puncturable pseudorandom sets and private information retrieval with near-optimal online bandwidth and time," in *Advances in Cryptology – CRYPTO 2021*, 2021, pp. 641–669. Available: https://doi.org/10.1007/978-3-030-84259-8_22

[221] H. Nissenbaum, "Contextual integrity up and down the data food chain," *Theoretical Inquiries in Law*, vol. 20, no. 1, pp. 221–256, 2019. Available: https://doi.org/10.1515/til-2019-0008

[222] K. Martin and H. Nissenbaum, "Privacy interests in public records: An empirical investigation," *Harvard Journal of Law & Technology*, vol. 31, no. 1, pp. 111–143, 2017. Available: https://jolt.law.harvard.edu/articles/pdf/v31/31HarvJLTech111.pdf

[223] National Institute of Standards and Technology, *Framework for Improving Critical Infrastructure Cybersecurity, version 1.1*, U.S. Department of Commerce. Gaithersburg, MD, April 2018. Available: https://doi.org/10.6028/NIST.CSWP.04162018

[224] National Institute of Standards and Technology, *Crosswalk from the NIST Privacy Framework Core to the Framework for Improving Critical Infrastructure Cybersecurity V1.1*, U.S. Department of Commerce. Gaithersburg, MD, January 2020. Available: https://www.nist.gov/privacy-framework/resource-repository/browse/crosswalks/cybersecurity-framework-crosswalk

[225] D. Huff, *How to Lie with Statistics*. New York: W. W. Norton & Company, 1954.

[226] T. F. Lunt, "Aggregation and inference: Facts and fallacies," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, May 1989, pp. 102–109. Available: https://doi.org/10.1109/SECPRI.1989.36284

[227] EdX, "Privacy policy," Wayback Machine, February 6, 2012. Accessed May 16, 2022. Available: https://web.archive.org/web/20120920014400/https://www.edx.org/privacy

[228] City of New York. (2005, May 19). Disclosure of Security Breach. Administrative Code, Title 10, Chapter 5. Available: https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=441345&GUID=8ECF99C5-52FA-48A2-B776-D5817E8D463A&Options=ID|Text|

[229] G. Slabodkin, "Defending DOD networks with a single security architecture," Defense Systems, July 19, 2013. Available: https://defensesystems.com/it-infrastructure/2013/07/defending-dod-networks-with-a-single-security-architecture/190685/

[230] *Department of Defense Privacy Program*, DoD Regulation 5400.11, Office of the Director, Administration and Management, DoD. Washington, DC, 2007. Available: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodm/540011r.pdf

[231] S. Groat, J. Tront, and R. Marchany, "Advancing the defense in depth model," in *2012 7th International Conference on System of Systems Engineering (SoSE)*, 2012, pp. 285–290. Available: https://doi.org/10.1109/SYSoSE.2012.6384127

[232] J. A. Whittaker, "No clear answers on monoculture issues," *IEEE Security & Privacy*, vol. 1, no. 6, pp. 18–19, 2003. Available: https://doi.org/10.1109/MSECP.2003.1266365

[233] City of New York, "Computer systems manager," Department of Information Technology and Telecommunications, Job Posting Notice, December 27, 2016. Available: https://www1.nyc.gov/assets/doitt/downloads/jobs/276549_dir_infrastructure_arch.pdf

[234] Chief Information Officer, U.S. Department of Defense, "DoDAF architecture framework version 2.02," August 2010. Available: https://dodcio.defense.gov/Library/DoD-Architecture-Framework/

[235] S. H. Dam, *DoD Architecture Framework: A Guide to Applying System Engineering to Develop Integrated, Executable Architectures*. Marshall, VA: SPEC, 2006.

[236] J. H. Saltzer and M. D. Schroeder, "The protection of information in computer systems," *Proceedings of the IEEE*, vol. 63, no. 9, pp. 1278–1308, 1975. Available: https://doi.org/10.1109/PROC.1975.9939

[237] H. Kelly, "A priest's phone location data outed his private life. It could happen to anyone." The Washington Post, July 22, 2021. Available: https://www.washingtonpost.com/technology/2021/07/22/data-phones-leaks-church/

[238] K. Dhondt, V. Le Pochat, A. Voulimeneas, W. Joosen, and S. Volckaert, "A run a day won't keep the hacker away: Inference attacks on endpoint privacy zones in fitness tracking social networks," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 2022, pp. 801–814. Available: https://doi.org/10.1145/3548606.3560616

[239] R. Hasan and M. Fritz, "Understanding utility and privacy of demographic data in education technology by causal analysis and adversarial-censoring," *Proceedings on Privacy Enhancing Technologies*, vol. 2022, no. 2, pp. 245–262, 2022. Available: https://doi.org/10.2478/popets-2022-0044

[240] D. Zhang, L. Yao, K. Chen, G. Long, and S. Wang, "Collective protection: Preventing sensitive inferences via integrative transformation," in *2019 IEEE International Conference on Data Mining (ICDM)*, 2019, pp. 1498–1503. Available: https://doi.org/10.1109/ICDM.2019.00197

[241] K. Crawford and J. Schultz, "Big data and due process: Toward a framework to redress predictive privacy harms," *Boston College Law Review*, vol. 55, no. 1, pp. 93–128, 2014. Available: https://core.ac.uk/reader/71464790

[242] K. Wuyts, L. Sion, and W. Joosen, "LINDDUN GO: A lightweight approach to privacy threat modeling," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 2020, pp. 302–309. Available: https://doi.org/10.1109/EuroSPW51379.2020.00047

[243] S. Shapiro *et al.*, "The PANOPTIC™ privacy threat model," Mitre Corporation, Bedford, MA, Technical Report, April 2024. Available: https://ptmworkshop.gitlab.io/static/media/PANOPTIC-MTR-v1-Publicly-Released.8cfab26fce008793a041.pdf

[244] "MITRE ATT&CK®," MITRE, 2024. Accessed on August 22, 2024. Available: https://attack.mitre.org/

[245] A. A. A. Zani, A. A. Norman, and N. A. Ghani, "A review of security awareness approaches: Towards achieving communal awareness," in *Cyber Influence and Cognitive Threats*, V. Benson and J. Mcalaney, Eds. Academic Press, 2020, pp. 97–127. Available: https://doi.org/10.1016/B978-0-12-819204-7.00006-3

[246] J. L. Kröger, L. Gellrich, S. Pape, S. R. Brause, and S. Ullrich, "Personal information inference from voice recordings: User awareness and privacy concerns," *Proceedings on Privacy Enhancing Technologies*, vol. 2022, no. 1, pp. 6–27, 2022. Available: https://doi.org/10.2478/popets-2022-0002

[247] M. Oates *et al.*, "Turtles, locks, and bathrooms: Understanding mental models of privacy through illustration," *Proceedings on Privacy Enhancing Technologies*, vol. 2018, no. 4, pp. 5–32, 2018. Available: https://doi.org/10.1515/popets-2018-0029

[248] K. Nissim and A. Wood, "Is privacy *Privacy*?" *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2128, p. 20170358, 2018. Available: https://doi.org/10.1098/rsta.2017.0358

[249] C. Graham, *Anonymisation: Managing Data Protection Risk Code of Practice*, Information Commissioner's Office. Cheshire, UK, November 2012. Available: https://ico.org.uk/media/1061/anonymisation-code.pdf

[250] C. Dwork, N. Kohli, and D. Mulligan, "Differential privacy in practice: Expose your epsilons!" *Journal of Privacy and Confidentiality*, vol. 9, no. 2, pp. 1–22, 2019. Available: https://doi.org/10.29012/jpc.689

[251] International Association of Privacy Professionals, "CIPP," 2022. Accessed May 2, 2022. Available: https://iapp.org/certify/cipp/

[252] F. Pasquale, "A rule of persons, not machines: The limits of legal automation," *George Washington Law Review*, vol. 87, no. 1, pp. 1–55, January 2019. Available: https://www.gwlr.org/wp-content/uploads/2019/01/87-Geo.-Wash.-L.-Rev.-1.pdf

[253] R. Crootof, ""Cyborg justice" and the risk of technological–legal lock-in," *Columbia Law Review*, vol. 119, no. 7, pp. 233–251, 2019. Available: https://www.jstor.org/stable/26960742

[254] *Information Security*, 44 U.S.C., Chapter 35, Subchapter II, United States Congress. Washington, DC, December 29, 2022. Available: https://uscode.house.gov/view.xhtml?path=/prelim@title44/chapter35/subchapter2&edition=prelim

[255] J. R. Biden, "Executive order 14028 – Improving the nation's cybersecurity," The White House, May 12, 2021. Available: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[256] J. R. Biden, *National Cybersecurity Strategy*, The White House. Washington, DC, March 2023. Available: https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf

[257] *Managing Information as a Strategic Resource*, Circular No. A-130, Office of Management and Budget. Washington, DC, 2016. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf

[258] Cybersecurity and Infrastructure Security Agency, ".Gov data," n.d. Accessed on October 10, 2023. Available: https://get.gov/about/data/

[259] Cybersecurity & Infrastructure Security Agency, "Federal civilian executive branch agencies list," n.d. Accessed November 22, 2023. Available: https://www.cisa.gov/news-events/directives/federal-civilian-executive-branch-agencies-list

[260] S. Kitterman, *Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1*, RFC 7208, Internet Engineering Task Force, April 2014. Available: https://datatracker.ietf.org/doc/html/rfc7208

[261] C. Wang *et al.*, "BreakSPF: How shared infrastructures magnify SPF vulnerabilities across the internet," in *Network and Distributed System Security (NDSS) Symposium*, 2024, pp. 1–18. Available: https://doi.org/10.14722/ndss.2024.23113

[262] Z. Durumeric *et al.*, "Neither snow nor rain nor MITM. . . An empirical analysis of email delivery security," in *Proceedings of the 2015 Internet Measurement Conference*, 2015, pp. 27–39. Available: https://doi.org/10.1145/2815675.2815695

[263] H. Hu and G. Wang, "End-to-end measurements of email spoofing attacks," in *27th USENIX Security Symposium (USENIX Security 18)*, August 2018, pp. 1095–1112. Available: https://www.usenix.org/conference/usenixsecurity18/presentation/hu

[264] S. Czybik, M. Horlboge, and K. Rieck, "Lazy gatekeepers: A large-scale study on SPF configuration in the wild," in *Proceedings of the 2023 ACM on Internet Measurement Conference (IMC '23)*, 2023, pp. 344–355. Available: https://doi.org/10.1145/3618257.3624827

[265] I. D. Foster, J. Larson, M. Masich, A. C. Snoeren, S. Savage, and K. Levchenko, "Security by any other name: On the effectiveness of provider based email security," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and*

Communications Security, 2015, pp. 450–464. Available: https://doi.org/10.1145/2810103.2813607

[266] M. S. Kucherawy and E. Zwicky, *Domain-based Message Authentication, Reporting, and Conformance (DMARC)*, RFC 7489, Internet Engineering Task Force, March 2015. Available: https://datatracker.ietf.org/doc/html/rfc7489

[267] New Jersey Cybersecurity & Communications Integration Cell (NJCCIC), "Phishing messages use email spoofing to impersonate NJ state office," Garden State Cyber Threat Highlight, April, 14 2023. Available: https://www.cyber.nj.gov/garden_state_cyber_threat_highlight/phishing-messages-use-email-spoofing-to-impersonate-nj-state-office

[268] R. A. Grimes. *How to Prevent 81% of Phishing Attacks from Sailing Right into Your Inbox with DMARC*. (2019). Accessed Aug. 11, 2022. Available: https://info.knowbe4.com/dmarc-spf-dkim-webinar

[269] B. Blechschmidt and B. Stock, "Extended hell(o): A comprehensive large-scale study on email confidentiality and integrity mechanisms in the wild," in *32nd USENIX Security Symposium (USENIX Security 23)*, August 2023, pp. 4895–4912. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/blechschmidt

[270] O. Hureau, J. Bayer, A. Duda, and M. Korczyński, "Spoofed emails: An analysis of the issues hindering a larger deployment of DMARC," in *Passive and Active Measurement*, vol. 14537 (LNCS), P. Richter, V. Bajpai, and E. Carisimo, Eds. Cham: Springer Nature Switzerland, 2024, pp. 232–261. Available: https://doi.org/10.1007/978-3-031-56249-5_10

[271] H. Hu, P. Peng, and G. Wang, "Towards understanding the adoption of anti-spoofing protocols in email systems," in *2018 IEEE Cybersecurity Development (SecDev)*, 2018, pp. 94–101. Available: https://doi.org/10.1109/SecDev.2018.00020

[272] B. Du, C. Testart, R. Fontugne, G. Akiwate, A. C. Snoeren, and k. claffy, "Mind your MANRS: Measuring the MANRS ecosystem," in *Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)*, 2022, pp. 716–729. Available: https://doi.org/10.1145/3517745.3561419

[273] P. Hoffman, *SMTP Service Extension for Secure SMTP over Transport Layer Security*, RFC 3207, Internet Engineering Task Force, February 2002. Available: https://datatracker.ietf.org/doc/html/rfc3207

[274] D. Tatang, R. Flume, and T. Holz, "A first large-scale analysis on usage of MTA-STS," in *Detection of Intrusions and Malware, and Vulnerability Assessment*, vol.

12756 (LNCS), L. Bilge, L. Cavallaro, G. Pellegrino, and N. Neves, Eds. Cham: Springer International Publishing, 2021, pp. 361–370. Available: https://doi.org/10.1007/978-3-030-80825-9_18

[275] I. Dolnák and J. Litvak, "Introduction to HTTP security headers and implementation of HTTP strict transport security (HSTS) header for HTTPS enforcing," in *15th International Conference on Emerging eLearning Technologies and Applications*, 2017. Available: https://doi.org/10.1109/ICETA.2017.8102478

[276] R. Oppliger, *SSL and TLS: Theory and Practice*, 3rd ed. Norwood, MA: Artech House Publishers, 2023. Available: https://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=30747334

[277] P. Patila, P. Narayankar, D. G. Narayan, and S. Meena, "A comprehensive evaluation of cryptographic algorithms: DES, 3DES, AES, RSA and Blowfish," in *Proceedings of the 2015 International Conference on Information Security & Privacy (ICISP2015)*, 2016, pp. 617–624. Available: https://doi.org/10.1016/j.procs.2016.02.108

[278] National Institute of Standards and Technology, "Update to current use and deprecation of TDEA," July 11, 2017. Available: https://csrc.nist.gov/News/2017/Update-to-Current-Use-and-Deprecation-of-TDEA

[279] K. A. McKay and D. A. Cooper, *Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations*, NIST SP 800-52, National Institute of Standards and Technology, U.S. Department of Commerce. Gaithersburg, MD, USA, August 2019. Available: https://doi.org/10.6028/NIST.SP.800-52r2

[280] P. Jindal and B. Singh, "RC4 encryption – a literature survey," in *Proceedings of the 2014 International Conference on Information Communication Technologies (ICICT2014)*, 2015, pp. 697–705. Available: https://doi.org/10.1016/j.procs.2015.02.129

[281] C. M. Chernick, C. Edington III, M. J. Fanto, and R. Rosenthal, *Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations*, NIST SP 800-52, National Institute of Standards and Technology. Gaithersburg, MD, USA, 2005. Available: https://doi.org/10.6028/nist.sp.800-52

[282] T. Polk, K. McKay, and S. Chokhani, *Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations*, NIST SP 800-52, National Institute of Standards and Technology. Gaithersburg, MD, USA, 2014. Available: https://doi.org/10.6028/nist.sp.800-52r1

[283] A. Popov, *Prohibiting RC4 Cipher Suites*, RFC 7465, Internet Engineering Task Force, February 2015. Available: https://datatracker.ietf.org/doc/html/rfc7465

[284] Cybersecurity & Infrastructure Security Agency, "cisagov / trustymail," June 8, 2023. Available: https://github.com/cisagov/trustymail

[285] T. L. Glade, "Comparing TLS and certificate configurations of government and non-government websites," M.S. thesis, Naval Postgraduate School, Monterey, CA, June 2024. Available: https://hdl.handle.net/10945/73127

[286] J. B. Kruskal, "An overview of sequence comparison," in *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison* (The David Hume Series on Philosophy and Cognitive Science Reissues), D. Sankoff and J. B. Kruskal, Eds. Stanford, CA: CSLI Publications, 1999, pp. 1–44. Original work published 1983.

[287] N. Teodoro, L. Gonçalves, and C. Serrão, "NIST cybersecurity framework compliance: A generic model for dynamic assessment and predictive requirements," in *2015 IEEE Trustcom/BigDataSE/ISPA*, 2015, vol. 1, pp. 418–425. Available: https://doi.org/10.1109/Trustcom.2015.402

[288] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979. Available: https://www.jstor.org/stable/4615733

[289] D. C. Howell, *Statistical Methods for Psychology*, 7th ed. Belmont, CA: Wadsworth Cengage Learning, 2010.

[290] J. L. Devore, *Probability and Statistics*, 6th ed. Belmont, CA: Thomson Brooks/Cole, 2004.

[291] M. L. McHugh, "The chi-square test of independence," *Biochemia Medica*, vol. 23, no. 2, pp. 143–149, June 2023. Available: https://doi.org/10.11613/BM.2013.018

[292] D. R. Bradley, T. D. Bradley, S. G. McGrath, and S. D. Cutcomb, "Type I error rate of the chi-square test of independence in $R \times C$ tables that have small expected frequencies," *Psychological Bulletin*, vol. 86, no. 6, pp. 1290–1297, 1979. Available: https://doi.org/10.1037/0033-2909.86.6.1290

[293] N. Pattnaik, S. Li, and J. R. C. Nurse, "A survey of user perspectives on security and privacy in a home networking environment," *ACM Computing Surveys (CSUR)*, vol. 55, no. 9, 2023. Available: https://doi.org/10.1145/3558095

[294] B. Stock, G. Pellegrino, F. Li, M. Backes, and C. Rossow, "Didn't you hear me? — Towards more successful web vulnerability notifications," in *Network and Distributed System Security (NDSS) Symposium*, 2018, pp. 1–15. Available: https://doi.org/10.14722/ndss.2018.23171

[295] M. Maass *et al.*, "Effective notification campaigns on the web: A matter of trust, framing, and support," in *30th USENIX Security Symposium (USENIX Security 21)*, August 2021, pp. 2489–2506. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/maass

[296] L. Qin, L. Chen, D. Li, H. Ye, and Y. Wang, "Understanding route origin validation (ROV) deployment in the real world and why MANRS action 1 is not followed," in *Network and Distributed System Security (NDSS) Symposium*, 2024, pp. 1–17. Available: https://doi.org/10.14722/ndss.2024.24214

[297] D. Poddebniak, F. Ising, H. Böck, and S. Schinzel, "Why TLS is better without STARTTLS: A security analysis of STARTTLS in the email context," in *30th USENIX Security Symposium (USENIX Security 21)*, August 2021, pp. 4365–4382. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/poddebniak

[298] S. De los Santos and J. Torres, "Analysing HSTS and HPKP implementation in both browsers and servers," *IET Information Security*, vol. 12, no. 4, pp. 275–284, 2018. Available: https://doi.org/10.1049/iet-ifs.2017.0030

[299] N. Samarasinghe and M. Mannan, "Another look at TLS ecosystems in networked devices vs. web servers," *Computers & Security*, vol. 80, pp. 1–13, 2019. Available: https://doi.org/10.1016/j.cose.2018.09.001

[300] S. Singanamalla, E. H. B. Jang, R. Anderson, T. Kohno, and K. Heimerl, "Accept the risk and continue: Measuring the long tail of government `https` adoption," in *Proceedings of the ACM Internet Measurement Conference (IMC '20)*, 2020, pp. 577–597. Available: https://doi.org/10.1145/3419394.3423645

[301] L. Segal, "Independence from political influence—A shaky shield: A study of ten inspectors general," *Public Integrity*, vol. 12, no. 4, pp. 297–314, 2010. Available: https://doi.org/10.2753/PIN1099-9922120401

[302] C. Spearman, "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 15, no. 1, pp. 72–101, 1904. Available: https://doi.org/10.2307/1412159

[303] R. D. Mason and D. A. Lind, *Statistical Techniques in Business and Economics*, 8th ed. Burr Ridge, IL: Irwin, 1993.

[304] A. M. Glenberg and M. E. Andrzejewski, *Learning from Data: An Introduction to Statistical Reasoning*, 3rd ed. New York: Lawrence Erlbaum Associates, Taylor & Francis Group, 2008.

[305] C. R. M. McKenzie, M. J. Liersch, and S. R. Finkelstein, "Recommendations implicit in policy defaults," *Psychological Science*, vol. 17, no. 5, pp. 414–420, 2006. Available: https://doi.org/10.1111/j.1467-9280.2006.01721.x

[306] C. Simoiu, "Secure by default: A behavioral approach to cyber security," Ph.D. dissertation, Stanford University, Department of Management Science and Engineering, Stanford, CA, August 2020. Available: https://purl.stanford.edu/fn897wh2511

[307] MSFTTracyP *et al.*, "Use DMARC to validate email," Microsoft, June 15, 2023. Available: https://learn.microsoft.com/en-us/microsoft-365/security/office-365-security/email-authentication-dmarc-configure?view=o365-worldwide

[308] MSFTTracyP, v-ratulach, chrisda, Dansimp, and Ratulch, "Set up SPF to help prevent spoofing," Microsoft, June 15, 2023. Available: https://learn.microsoft.com/en-us/microsoft-365/security/office-365-security/email-authentication-spf-configure?view=o365-worldwide

[309] E. Lam *et al.*, "Poster: On the (in)security of government web and mail infrastructure," in *Network and Distributed System Security (NDSS) Symposium*, 2024, pp. 1–3. Available: https://www.ndss-symposium.org/wp-content/uploads/ndss24-posters-42.pdf

[310] A. K. Fetterman, S. Curtis, J. Carre, and K. Sassenberg, "On the willingness to admit wrongness: Validation of a new measure and an exploration of its correlates," *Personality and Individual Differences*, vol. 138, pp. 193–202, 2019. Available: https://doi.org/10.1016/j.paid.2018.10.002

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
   Fort Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California