Federal Zero Trust Data Security Guide

OCTOBER 2024

TABLE OF CONTENTS

TABLE OF CONTENTS

FIGURES

TABLES

AUTHORITY

This document was developed by the **Zero Trust (ZT)** Data Security Working Group in furtherance of its directive under the **Office of Management and Budget (OMB)** Memorandum M-22-09, *Moving the U.S. Government Towards Zero Trust Cybersecurity Principles*.[1] The Working Group is a joint committee comprised of members from the Federal **Chief Data Officer (CDO)** Council and **Chief Information Security Officer (CISO)** Council, as well as other Federal stakeholders. OMB M-22-09 charges the Working Group with: (1) developing a guide for Federal agencies that addresses how existing Federal information categorization schemes can support effective data categorization in a security context; (2) developing enterprise-specific data categories that are not addressed by existing Federal categories; (3) identifying members who will act as leads, or designate leads within their agencies, to convene a **community of practice (CoP)** that can assist agencies in tackling specific areas of focus; and (4) identifying and supporting pilots of emerging approaches and best practices among agencies. This document is intended to assist agencies on their ZT journey as they continue to implement ZT principles year-to-year and across Administrations.

Disclaimer

This document is intended to provide best practices to help agencies consider data security issues that are relevant to the implementation of ZT requirements. This document does not establish or modify any requirements in law, regulation, or policy. The best practices presented in this document are intended to be consistent with Federal policies and laws, including, but not limited to: **Executive Order (EO)** 14028 on *Improving the Nation's Cybersecurity*;[2] EO 13744 on *Making Open and Machine Readable the New Default for Government Information*;[3] Foundations for Evidence-Based Policymaking Act (Pub. L. 115–435); and OMB M-19-18, *Federal Data Strategy — A Framework for Consistency*.[4] Nothing in this document may be construed to replace or supersede any authorities, policies, guidance, standards, or other articles made mandatory and binding for Federal agencies under statute. Agencies should consult with their counsel, policy officials, and relevant stakeholders to determine the appropriate process for implementing ZT data security requirements.

This document is non-binding and is a reference for Federal agencies. While nongovernmental organizations may use this document on a voluntary basis and this document is not subject to copyright regulations, attribution would be appreciated by the Federal **Chief Information Officer (CIO)** Council, CISO Council, and CDO Council.

[1] OMB M-22-09: *Moving the U.S. Government Towards Zero Trust Cybersecurity Principles*, https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf

[2] EO 14028 on *Improving the Nation's Cybersecurity*, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[3] EO 13744 on *Making Open and Machine Readable the New Default for Government Information*, https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

[4] OMB M-19-18: *Federal Data Strategy – A Framework for Consistency*, https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf

This document focuses on the Data Pillar of the **Cybersecurity and Infrastructure Security Agency's (CISA) ZT Maturity Model (ZTMM)**[5] and is not intended to cover every aspect of ZT. Certain commercial entities, equipment, or materials may be referenced in this document to adequately describe an experimental procedure, theoretical application, or concept. Such reference is not intended to imply recommendation or endorsement by the Working Group, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

ACKNOWLEDGEMENTS

This document was developed by CDOs, CISOs, **Senior Agency Officials for Privacy (SAOPs)**, and representatives from over 30 Federal departments, agencies, bureaus, and entities, in consultation with other stakeholders.

[5] CISA ZTMM, https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf

CHAPTER 1: INTRODUCTION

1.1: DATA MANAGEMENT IS CRITICAL TO MAKING ZERO TRUST A REALITY

Our vision is to establish a resilient and secure digital environment where users, assets, and access to resources are continuously validated and verified.

The cyber risk landscape is continuously evolving — and our adversaries are evolving along with it. The United States is facing unprecedented threats as malicious actors advance their tactics and unlock new ways to attack our systems, including using emerging technologies, such as **artificial intelligence (AI)**, to launch increasingly sophisticated cyber campaigns.

To counter these threats, agencies are making Federal systems more defensible by employing ZT principles — which means trust is never implicitly granted and must be continually validated.

ZT moves away from the traditional approach of protecting the network perimeter — a "castle and moat" model as seen in Figure 1 — to instead assume that a network may be compromised at any time, anywhere, and by anyone. Through the ZT lens, we focus on **securing the data itself, rather than the perimeter protecting it.** This concept is known as "ZT data security."

FIGURE 1: Traditional Security vs. Zero Trust

Traditional Security

Traditional network perimeter-based security, with its assumption of implicit trust inside the perimeter, has failed to protect enterprise assets. This belief that everything is safe and can be trusted once inside actually paves the way for attackers to cause chaos through unimpeded lateral movement.

Zero Trust

Zero trust assumes that all networks — enterprise-owned or not — are untrusted and that an attacker is present in the environment. It denies default access to data and workload, continually authenticates and authorizes each access request, and monitors and analyzes the risks to the assets.

ZT data security is not a tool that you can buy off the shelf and achieve overnight — it is a multi-year strategy that requires adaptability, investment, and collaboration. This guide aims to help agencies on their journey to ZT data security by being practical and operational, yet flexible enough for practitioners to adjust to meet their agency's unique needs and mission requirements. The intended audience for this guide includes practitioners charged with securing data, including system owners, data management practitioners and stewards, system administrators, and cybersecurity engineers.

This guide breaks the ZT data security journey into three chapters:

1. **Define the Data:** This chapter helps readers find and identify the totality of their data landscape, learn how to accurately categorize and handle data, and define the sensitivity and criticality of data.
2. **Secure the Data:** This chapter focuses on implementing the appropriate security monitoring and controls, such as encryption, for data and incorporating risk management and **identity, credential and access management (ICAM)**. It also explains considerations for privacy and compliance.
3. **Manage the Data:**[6] The goal of this chapter will be to ensure that data security practices are aligned with and embedded in **data lifecycle management (DLM)**. It will discuss how readers can equip their team with the necessary skills and adapt their approach to address emerging technologies.

Implementing a holistic security approach like ZT is challenging in any environment, especially where components may operate independently. However, we must remember that cybersecurity is a team sport and — with demonstrated senior leader buy-in and stakeholder collaboration — we can overcome the challenge. Looking ahead, we must build and sustain a cooperative relationship between data management and cybersecurity teams across Government. By embracing innovation, cultivating partnerships, and acting swiftly, we can protect our data from adversaries.

[6] Placeholder notice — We've included a placeholder for this chapter as we continue to develop and refine content. Our priority is to share insights at the speed of need, and we hope that readers can leverage the aforementioned chapters to begin the process of defining and securing their data.

1.2: CONNECTING THE DOTS BETWEEN ZERO TRUST AND DATA

Data security is at the heart of Zero Trust.

Protecting sensitive data assets is at the heart of the ZT model. As such, data management practices must be **secure by design**, where the security of data is treated as a core requirement from the beginning of and throughout its lifecycle — security cannot be a mere aspiration or afterthought. The ZT model forces security upstream in DLM, starting at the logical design stage of the data as depicted in Figure 2.

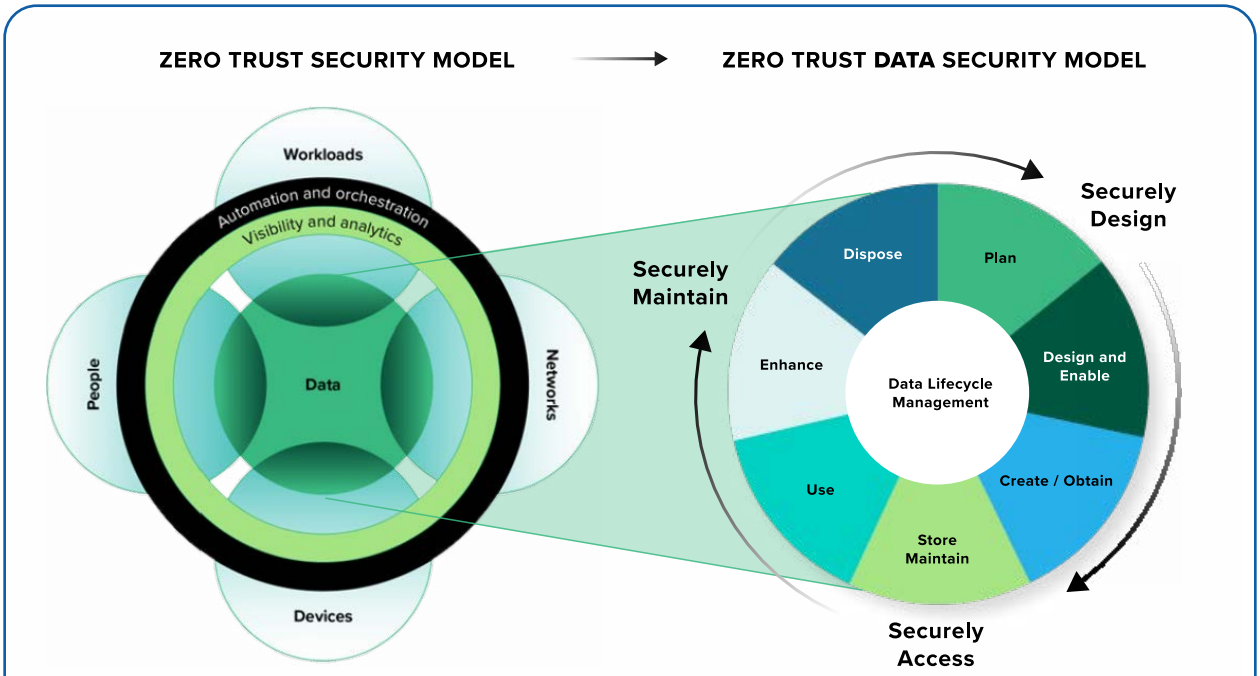
Establishing security as a core design criterion from the outset can also help facilitate seamless transition of data into other environments. This is particularly useful as agencies are increasingly moving their data to cloud-based, shared environments.

Data as the New Perimeter

"[We must] … extend the protect surface by defining data down to the cellular level as the new perimeter and adopting dynamic data tagging, labeling, and encryption technology that empowers data stewards and consumers and assures access from anywhere, anytime."

— Department of the Air Force *Zero Trust Strategy*, DAF Zero Trust Strategy v1.0 - SAF/C, https://www.safcn.af.mil/Portals/64/Documents/Strategy/DAF%20Zero%20Trust%20Strategy%20v1.0%20(002).pdf

FIGURE 2: Secure Data Lifecycle Management Is at the Heart of the Zero Trust Model



Source: Zero Trust Security Model, The Business of Zero Trust Security, Forrester Research, Inc., July 11, 2023

1.3: ZERO TRUST DATA SECURITY PRINCIPLES

Practitioners can reference the principles in Table 1 to inform decision-making and ensure alignment with their agency's strategy, mission, and values.

TABLE 1: Zero Trust Data Security Principles

Principle	Why?
Adopt a data-centric view	Data is everywhere and exists in different formats with varying levels of sensitivity and value. Protecting critical data requires visibility, analytics, and automation across the entire digital ecosystem.
Implement standardized least privilege and strictly enforce access control	The bedrock principles of ZT are that all entities are untrusted, least privilege access is enforced, and comprehensive security monitoring is implemented.
Promote data resiliency and integrity	The value of data is maximized when it is available, accessible, and trustworthy. The Federal government relies on quality data to conduct business and deliver services to the public.
Integrate data and security literacy	Data and security practitioners must understand each other's nomenclature to effectively safeguard their agencies' data and enable appropriate use.
The impact of data security must be measurable and actionable	Meaningful analytics that produce actionable insights can help to prevent breaches and reduce the impact of breaches when they do occur.
Data security is risk-informed throughout the data lifecycle	Each stage of the data lifecycle has specific security requirements. Security controls must address risks **to** the data, **from** the data, and **in** the data.
Balance priorities — make the most with what you have	ZT principles will shape existing practices, processes, and perimeters. As practitioners understand these changes, they can assess whether their current cyber infrastructure meets these evolved needs.

DID YOU KNOW?

The DAMA Guide to the **Data Management Body of Knowledge (DAMA-DMBOK2)** is a well-known reference guide for data management professionals, and has been used by practitioners at Federal agencies.

The DAMA-DMBOK2 defines the goals of data security as:

- Enable appropriate, and prevent inappropriate, access to enterprise data assets.
- Understand and comply with all relevant regulations and policies for privacy, protection, and confidentiality.
- Ensure that the privacy and confidentiality needs of all stakeholders are enforced and audited.

1.4: APPLYING ZERO TRUST DATA SECURITY GUIDANCE

The journey to ZT data security will be complex and extensive, so practitioners should develop a ZT data security roadmap to help manage the task effectively. Practitioners should ensure the roadmap aligns with their agency's ZT implementation plan, as required by EO 14028[7] and OMB M-22-09.[8]

Practitioners should start by assessing their agency's current maturity level and identifying the desired future maturity level with a clear timeframe. Utilizing an existing maturity model, such as CISA's ZTMM,[9] provides a way to evaluate the agency's current state, identify the target state goals and metrics to measure progress toward those goals, and refine the roadmap as needed.

If a practitioner's agency doesn't have a published or fully mature ZT implementation plan, they may draw inspiration from other agencies' published plans. While this guide can be used as a starting point, practitioners are encouraged to leverage other existing roadmaps and resources. Remember — don't reinvent the wheel or create silos!

[7] EO 14028 on *Improving the Nation's Cybersecurity*, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[8] OMB M-22-09: *Moving the U.S. Government Toward Zero Trust Cybersecurity Principles*, https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf

[9] CISA ZTMM, https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf

To inform the drafting of their roadmap, practitioners should:

1. Assess the current maturity level of data management and security of their organization.
2. Consider the unique needs of their organization, including current and planned changes to cybersecurity posture, mission(s), and existing or anticipated organizational technology, strategic, business, and data management roadmaps and changes.
3. Understand the existing landscape of business initiatives, technology transformations, and security projects.
4. Document where existing data management and data security capabilities can be reused and where new capabilities are needed.
5. Identify and prioritize related initiatives such as **controlled unclassified information (CUI)** and privacy.
6. Set goals for future maturation and identify a timeframe to achieve those goals.

These steps may have already been performed in the agency's ZT implementation plan. If so, practitioners should verify the findings are up to date before proceeding.

1.4.1: UNDERSTAND THE STARTING POINT — FOCUS ON SMALL WINS AND BUILD OUT

A ZT data security roadmap should address the stages outlined in Table 2. Although Table 2 includes *suggested* actions and leads for each stage, practitioners should determine whether these are applicable to their organization.

TABLE 2: Stages of a Zero Trust Data Security Roadmap

Stage	What	Who
1. Discover, inventory, categorize, label, and map the data flows *See* Chapter 2 *to learn more*	• Establish the foundation and infrastructure for the ZT data security roadmap. • Find all the data across the agency. • Map the critical data flows between existing systems, processes, and data stores to discover the trust relationships and logical mission trust zones.	**Lead:** The agency's CDO or designated data steward(s) should lead this stage to develop an understanding of the current data landscape, data stewardship, categorization, and labeling in metadata. **Coordinate with:** Data governance body, SAOP, records retention officials, and other stakeholders early in the process to identify the steps in the roadmap
2. Conduct risk analysis *See* Chapter 3 *to learn more*	• Conduct a risk analysis to articulate data integrity issues that can cause financial, reputational, third-party, and supply chain losses. • Prioritize focus areas based on risk, impact, and achievability.	**Lead:** The agency's ZT or cybersecurity practitioners typically lead this and subsequent stages. **Coordinate with:** Data management and privacy practitioners and agency experts involved in ongoing risk analysis and threat modeling

Stage	What	Who
3. Align with the ZT architecture	• Review the agency's ZT implementation plan. • Set timelines and quick feedback mechanisms.	**Lead:** The agency's ZT or cybersecurity practitioners typically lead this and subsequent stages. **Coordinate with:** The team responsible for implementing ZT, such as the agency's ZT Program Office, and data management practitioners.
4. Design the controls and monitoring *See* Chapter 3 *to learn more*	• Identify tools to close data security gaps. • Implement data-centric security controls, monitoring, logging, and alerting. • Refactor the architectural, **Federal Information Security Modernization Act (FISMA)**, financial boundaries, and data standardization and centralization.	**Lead:** The agency's ZT or cybersecurity practitioners typically lead this and subsequent stages. **Coordinate with:** Data management and privacy practitioners and agency experts involved in ongoing risk analysis and threat modeling.
5. Embrace automation and orchestration	• Identify opportunities to integrate and orchestrate automation to manage the operational data.	**Lead:** The agency's CDO or designated data steward(s) should lead this stage. **Coordinate with:** Security and privacy practitioners and agency experts involved in automation (e.g., **Chief AI Officer [CAIO]**)

CHAPTER 2: DEFINE THE DATA

While agencies have been required to inventory their data assets for some time, a comprehensive ZT approach to data management requires going beyond and — in some ways — *redefining* what agencies may be accustomed to thinking of as "data." This chapter provides guidance on data management and governance practices essential to ZT data security. It leverages a rich, established body of work published by the CDO Council rooted in existing Federal statutes and **National Institute of Standards and Technology (NIST)** guidelines.

Agencies must first identify and locate their sensitive data assets, inventory those assets, establish data categories using norms aligned to their missions, and label those assets accordingly. This chapter focuses on how to:

Agency data management and privacy practitioners should lead the operationalization of the guidelines outlined in this chapter, working closely with cybersecurity practitioners to conduct or update their data inventory.

1. Identify and locate data assets to understand the entirety of an agency's sensitive data landscape.
2. Create a data asset catalog and inventory that aligns with the agency's mission and needs as the single source of truth for agency data assets.
3. Categorize the data assets so the appropriate security controls can be applied.
4. Label (i.e., mark or tag) the data sources and associated metadata to support automated protection.

2.1: LEGAL AND REGULATORY REQUIREMENTS FOR DATA INVENTORY AND CATEGORIZATION

Multiple laws, regulations, and Federal guidance already require the maintenance of an inventory of data assets. As such, agencies may be able to build on existing inventories to jumpstart their ZT data inventory, categorization, and labeling initiative. The list below provides examples of existing laws, regulations, and guidance that may require inventory-related efforts. Note that this list only includes examples in effect at the time of this guide's publication, and **it's the agency's responsibility to stay up-to-date with the latest legal requirements.** Examples include, but are not limited to:

- The Foundations for Evidence-Based Policy Making Act of 2018 ("Evidence Act," Pub. L. 115-435)[10]

[10] The Foundations for Evidence-Based Policy Making Act of 2018, https://www.congress.gov/bill/115th-congress/house-bill/4174

- OMB M-21-27, *Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans*[11]
- OMB M-19-18, *Federal Data Strategy — A Framework for Consistency*[12]
- OMB M-23-04, *Establishment of Standard Application Process Requirements on Recognized Statistical Agencies and Units*[13]
- The Paperwork Reduction Act of 1980 (Pub. L. 96–511)[14]
- The Privacy Act of 1974 (Pub. L. 93-579)[15]
- The E-Government Act of 2002 (Pub. L. 107-347)[16]
- FISMA 2014 (Pub. L. 113-283)[17]
- The **Health Insurance Portability and Accountability Act (HIPAA)** of 1996 (Pub. L. 104-191)[18]
- "Desirable Characteristics of Data Repositories for Federally Funded Research,"[19] published by the National Science and Technology Council
- CISA **Binding Operational Directive (BOD)** 18-02: Securing **High Value Assets (HVA)**[20]
- CISA ZTMM Version 2.0[21]
- "A Framework for Data Quality,"[22] published by the **Federal Committee on Statistical Methodology (FCSM)**

[11] OMB M-21-27: *Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans*, https://www.whitehouse.gov/wp-content/uploads/2021/06/M-21-27.pdf

[12] OMB M-19-18: *Federal Data Strategy — A Framework for Consistency*, https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf

[13] OMB M-23-04: *Establishment of Standard Application Process Requirements on Recognized Statistical Agencies and Units*, https://www.whitehouse.gov/wp-content/uploads/2022/12/M-23-04.pdf

[14] The Paperwork Reduction Act of 1980, https://www.congress.gov/96/statute/STATUTE-94/STATUTE-94-Pg2812.pdf

[15] The Privacy Act of 1974, https://www.congress.gov/93/statute/STATUTE-88/STATUTE-88-Pg1895.pdf

[16] The E-Government Act of 2002, https://www.congress.gov/bill/107th-congress/house-bill/2458

[17] FISMA 2014, https://www.congress.gov/bill/113th-congress/senate-bill/2521

[18] HIPAA 1996, https://www.congress.gov/bill/104th-congress/house-bill/3103/text

[19] Desirable Characteristics of Data Repositories for Federally Funded Research, https://www.whitehouse.gov/wp-content/uploads/2022/05/05-2022-Desirable-Characteristics-of-Data-Repositories.pdf

[20] CISA BOD 18-02: Securing High Value Assets, https://www.cisa.gov/news-events/directives/bod-18-02-securing-high-value-assets

[21] CISA ZTMM, https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf

[22] A Framework for Data Quality, https://nces.ed.gov/FCSM/pdf/FCSM.20.04_A_Framework_for_Data_Quality.pdf

2.2: DEFINING THE DATA ASSETS

An agency's data landscape may seem daunting and uncharted due to the sheer volume and diversity of types, formats, locations, and use cases. It can encompass **personally identifiable information (PII)**[23] stored in databases and files on endpoints, digital conversations saved in cloud environments, structured and unstructured data in data lakes and file stores, and more.

Since data is so vast and ubiquitous, agencies need a shared understanding of what data assets are in order to identify and protect them.

The Evidence Act defines data as "recorded information, regardless of form or the media on which the data is recorded" and a data asset as "a collection of data elements or data sets that may be grouped together."[24] The CISA ZTMM defines data as "all structured and unstructured files and fragments that reside or have resided in Federal systems, devices, networks, applications, databases, infrastructure, and backups (including on-premises and virtual environments) as well as the associated metadata."[25] Generally, Federal CDOs have taken the view that a data inventory is not just a "collection of data" or "recorded information," but an **intentional, meaningful categorization of data that serves a mission purpose and brings value to the agency.**

2.3: COMMON STARTING POINT TO IDENTIFY SENSITIVE DATA ASSETS AND BEGIN THE DATA CATALOG

As mentioned earlier, in many cases, agencies won't need to start their inventory from scratch— they may be able to build upon existing lists of data assets from other efforts highlighted above. Agencies must aim to organize these assets around a meaningful taxonomy based on the agency's mission and business areas. This will serve as the foundation for the data catalog and ensure that the data is organized in a way that is most useful for the agency.

While each agency will develop its unique taxonomy of data assets, a good starting point is to identify the mission-based information types. We recommend referencing NIST **Special Publication (SP)** 800-600 Rev. 2, *Guide for Mapping Types of Information and Systems to Security Categories*.[26] Data practitioners can also collaborate with the agency enterprise architects to use a Business Reference Model of Federal Enterprise Architecture,[27] if applicable.

---

[23] Per OMB Circular A-130, PII is defined as "information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual." See https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf.

[24] The Foundations for Evidence-Based Policy Making Act of 2018, https://www.congress.gov/bill/115th-congress/house-bill/4174

[25] CISA ZTMM, https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf

[26] Guide for Mapping Types of Information and Systems to Security Categories, https://doi.org/10.6028/NIST.SP.800-60r2.iwd

[27] Business Reference Model of Federal Enterprise Architecture, https://obamawhitehouse.archives.gov/omb/e-gov/FEA

2.4: CREATING THE DATA INVENTORY

The Evidence Act requires agencies to "develop and maintain a comprehensive data inventory that accounts for all data assets created by, collected by, under the control or direction of, or maintained by the agency," to the extent practicable.[28] The CDO Council has been leading the charge in identifying best practices for data inventories.

In April 2022, the CDO Council released the Enterprise Data Inventories Report,[29] which states:

> *"...a data inventory is foundational to any formal data management program. Data inventories enable stakeholders to efficiently find, access, and use data assets. Inventories are also indispensable to managers who need to evaluate the extent to which the organization's data helps meet mission goals or can be shared with other agencies to meet their missions."*

The CISA ZTMM echoes that a comprehensive data inventory is a factor in improving the ZT maturity of Federal agencies. Increasing the automation of the inventory can further boost an agency's ZT maturity from "traditional" to "optimal." Therefore, many agencies that have started their ZT journey may consider prioritizing the data inventory as an immediate action. Agencies can approach the inventory through manual or automated creation as outlined in Table 3 below.

TABLE 3: Approaches To Creating the Data Inventory

Manual Creation	Automated Creation
Agencies may find that starting with HVAs and FISMA High Impact systems makes the most sense. This is because agencies are likely to have already established these assets and systems in accordance with Federal laws and policies.[30] Agencies should capture relevant data asset taxonomy and metadata in commonly available tools (e.g., Excel). It is recommended that agencies use standard tools as a starting point until they acquire modern data catalog software.	Some agencies have invested in **machine learning data catalog (MLDC)** tools, which utilize advanced algorithms and techniques to automate capabilities such as data discovery, metadata extraction, and other data management activities. Many agencies may have existing technologies and data security tools within their **information technology (IT)** environment with sensitive data discovery and classification capabilities. Agencies can use the insights generated from these tools' data discovery capabilities to help develop a data inventory.

For more details about the steps for creating a data inventory, see *Appendix A: Data Inventory* in the companion document.

[28] The Foundations for Evidence-Based Policy Making Act of 2018, https://www.congress.gov/bill/115th-congress/house-bill/4174

[29] CDO Council Data Inventory Report, https://resources.data.gov/assets/documents/CDOC_Data_Inventory_Report_Final.pdf

[30] See OMB M-19-03: *Strengthening the Cybersecurity of Federal Agencies by enhancing the High Value Asset Program*, https://www.whitehouse.gov/wp-content/uploads/2018/12/M-19-03.pdf and FISMA 2014

2.5: STANDARDS FOR THE DATA INVENTORY

Since standards for data catalogs and inventories are evolving, agencies should be vigilant about development in this space. For example, an updated data catalog standard for the United States, DCAT-US v3.0 Schema,[31] is currently being developed. It builds on the DCAT standard, DCAT-US v1.1,[32] which Data.gov has used for over a decade. This standard, in addition to the **DATA Act Information Model Schema (DAIMS)**,[33] can help practitioners set internal data catalog standards for their agency. Additionally, NIEMOpen[34] — while not a standard for data inventories — can be used to help establish common vocabulary for easier exchange of information.

Security practitioners are encouraged to collaborate with their data management counterparts to gain a foundational understanding of the data inventory and related standards. Ultimately, security practitioners should learn to apply the data management discipline to manage security-relevant, non-business data, such as logs.

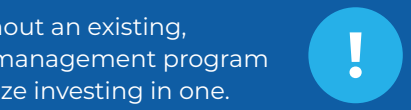
Agencies can find additional resources at https://resources.data.gov/standards/catalog/, a repository of data standards.

2.6: ROLES AND RESPONSIBILITIES

Before diving into the steps of data categorization, it is helpful to understand the roles and responsibilities of various parties involved in the process.

In agencies with mature data management programs, the data governance team typically develops the data catalog and data inventory. These agencies have established roles, such as business data steward and technical data steward, to create and maintain their data inventory, business glossary, and metadata library. Data stewards are great resources for identifying business risks. They should collaborate with their agency's security team to do risk-based categorization and labeling of their data assets. Data stewards should also collaborate with the agency's legal, records management, and privacy officials to establish the data inventory.

Agencies without an existing, mature data management program should prioritize investing in one.



While the implementation of CUI standards is underway in many agencies, a new role of information steward is also emerging. An information steward is similar to a data steward except that they steward unstructured data assets. They are the business experts likely to be best positioned to understand the risks of unstructured data assets and assign appropriate categories.

[31] DCAT-US v3.0 Schema, https://github.com/DOI-DO/dcat-us

[32] DCAT-US v1.1, https://resources.data.gov/resources/dcat-us/

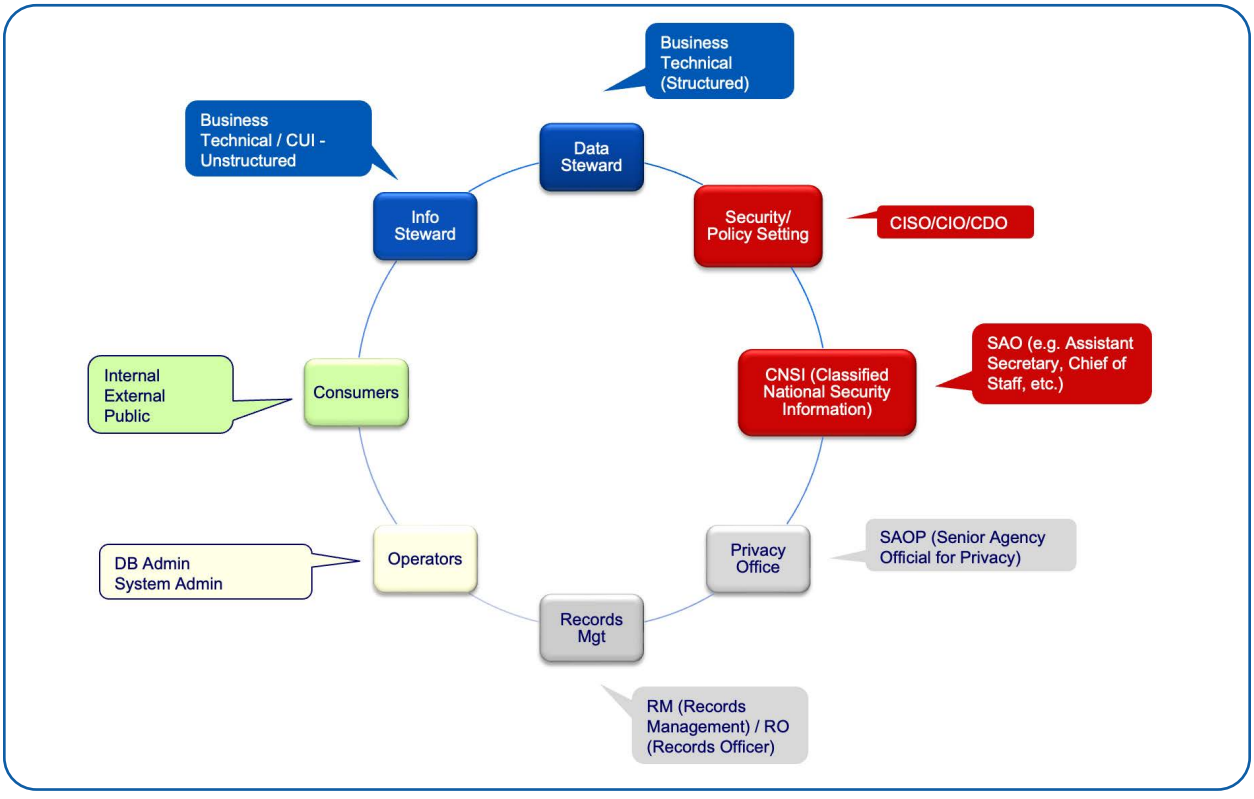
[33] DATA Act Information Model Schema (DAIMS), https://resources.data.gov/standards/catalog/daims/

[34] NIEMOpen, https://www.niem.gov/

Additionally, there are other key players in ZT data security as outlined in Figure 3. The following roles, bodies, or their designee(s) should likely be consulted in creating the data inventory and categorization:

- Agency Data Governance Body or Board
- Agency Records Officer or Senior Agency Official for Records Management
- **Agency Statistical Official (ASO)**
- CAIO
- CDO
- CIO
- CISO
- SAOP
- Senior Agency Official for CUI
- **Senior Agency Official for Geospatial Information (SAOGI)**

FIGURE 3: Zero Trust Data Security Players



Refer to *Appendix B: Data Stewardship* in the companion document for additional details about the role of the agency stewards.

2.7: CATEGORIZATION OF DATA ASSETS BASED ON IMPACT

The loss or misuse of data assets have different impacts on the mission of the Federal government. Understanding the impact of data's loss or misuse is the first step to understanding how data should be categorized and secured.

Agencies can refer to several well-known frameworks for categorizing data assets based on impact:

- **Federal Information Processing Standards (FIPS) Publication 199, *Standards for Security Categorization of Federal Information and Information Systems*:**[35] Agencies should already be familiar with applying the standards set by FIPS 199 for semi-structured data assets, such as databases or XML-based data assets with underlying structure. Agencies can adapt this framework to apply to broader datasets, focusing on categorization of all information types housed throughout the system.
- **CUI:** Established by EO 13556 on *Controlled Unclassified Information*,[36] the CUI program is a great starting point for categorizing and labeling unstructured data — including emails, documents, spreadsheets, and text files — that are typically stored in a file system repository. Some agencies have found applying a common CUI labeling schema to structured or semi-structured data to be a streamlined approach.
- **NIST SP 800-60 Vol. 2 Rev. 1, *Guide for Mapping Types of Information and Information Systems to Security Categories: Appendices*:**[37] This publication contains security categorization recommendations and rationale for mission-based, management, and support information types. It includes provisional impact level recommendations for various administrative, management, and service information.

[35] FIPS 199, Standards for Security Categorization of Federal Information and Information Systems, https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.199.pdf

[36] EO 13556 on *Controlled Unclassified Information*, https://obamawhitehouse.archives.gov/the-press-office/2010/11/04/executive-order-13556-controlled-unclassified-information

[37] NIST SP 800-60, *Guide for Mapping Types of Information and Information Systems to Security Categories: Appendices*, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v2r1.pdf

2.8 DATA CATEGORIES AND LABELS

Agencies must develop or adopt methodologies to label their data assets. We recommend using the CUI sensitivity types as a starting point for labeling sensitive data, whether structured, semi-structured, or unstructured. Labeling should be done at the document level only and not at the data attribute level. Agencies can also use records management file plans that identify and categorize data according to the **National Archives and Records Administration (NARA)** Federal Records Schedule. For example, most Federal contracts are managed using **General Records Schedules (GRS)** 1.[38]

Label Deployment: Labels should not be deployed to the whole agency all at once. The best practice is to deploy labels in an incremental fashion when a business need is identified.

- **Initial Labels:** Labels should strive to address the largest percentage of data categories within an agency (e.g., privacy, financial).
- **Label Scoping:** Some labels may only apply to specific agencies (via Active Directory Groups). Some organizations dealing with highly sensitive data may warrant special protection labels.
- **Methods:** Technologies enable manual and automated labeling. There can be a need for both methods, as well as challenges associated with each approach.
    - **Manual labeling** can be useful in instances where automated categorization and labeling is difficult to achieve accuracy due to the type of content or data. Manual labeling is challenging to scale, can have adverse impacts to employee workflows, and is prone to human error and manipulation.
    - **Automated labeling** capabilities are improving across tools today, though agencies must test this functionality to validate the accuracy of the categorization and labels. Automated labeling is increasingly necessary due to the growing volumes of data created today and in the future.

[38] General Records Schedules, https://www.archives.gov/records-mgmt/grs

CHAPTER 3: SECURE THE DATA

This chapter provides guidance on managing common data security risks. For the purpose of this guide, data security risks refer to potential threats and/or vulnerabilities that could compromise the confidentiality, integrity, and availability of an agency's sensitive data. These risks are amplified by the expanding capabilities of digital networks, AI, algorithms, and computational methods that enable data to be easily collected, combined, manipulated, and shared. Risk elements are outlined in Table 4.

TABLE 4: Risk Elements

Risk Elements	Data-Centric Guidance
Threat	A threat is any circumstance or event caused via unauthorized access, destruction, disclosure, or modification of data that the information system maintains. In other words, most threats target the data. Threats can be adversarial and non-adversarial.
Vulnerability	This is a weakness in an information system, system security procedures, or internal controls. While weaknesses are typically in the security controls that protect the data, they can also be in the data itself (e.g., lack of integrity, poor quality).
Likelihood	The probability of an adverse event or security incident depends on a combination of factors, such as: 1. Likelihood the event will occur. 2. Initiation by an adversary. 3. Likelihood that the event will result in adverse impacts. From the data perspective, the likelihood of an adverse event or security incident depends on the value of the data to the organization and its adversaries. An adversary is more likely to apply resources to higher value assets and data.[39]
Impact	The impact to an organization of an adverse security incident is based on the value and sensitivity of the data to stakeholders (e.g., the organization and its customers). See *Chapter 2* for more information on assessing impact.

[39] CISA's HVA Program Management Office connects the authorities of OMB and CISA to identify the most valuable systems and harmonize a government-wide approach to protect HVA system functions and the information they contain. For more information, visit https://www.cisa.gov/resources-tools/programs/high-value-asset-program-management-office.

3.1: A BRIEF ANATOMY OF DATA SECURITY RISKS

Data security risks can be unintentional or intentional and stem from sources such as:

- External actors.
- Insider threats.
- Trusted third parties.
- Human/machine errors.
- Natural disasters.

Additionally, data security risks can stem from the data itself, such as issues related to data integrity, quality, and lifecycle management.

These risks have the potential to harm the national security and economic interests of the United States, so it's imperative that practitioners understand their roles and responsibilities and are held accountable for managing information security risk.[40]

Consider the following representative examples of data security risk types outlined in Table 5.

TABLE 5: Data Security Risk Types

Risk Type	Representative Examples
Risks **TO** the data *Risks to the confidentiality, integrity, and/or availability of data from external and internal threat actors, or from an agency's own IT practices*	• **Cybersecurity threats:** Ransomware, data extortion, insider threats, data breaches, eavesdropping, person-in-the-middle attacks, or accidental misuse of data • **Data transmission risks:** Insecure communication channels for data in transit, such as lack of secure encryption, improper use or configuration of encryption modules, hardware, ciphers, etc. • **Data storage risks:** Failure to properly encrypt data at rest, such as insecure or improper usage or configuration of encryption algorithms, cross origin sharing policies, or improper encryption key management • **Data resiliency risks:** Incomplete or untested backups, lack of redundancy, or loss in the availability and integrity of data • **DLM risks:** Improper handling of data throughout its lifecycle, including its initiation, development, implementation, operation, maintenance, and disposal • **Storage failure:** Data loss or corruption due to physical or logical failures of storage systems or devices • **Incomplete erasure:** Unauthorized use of residual data that is not fully erased from storage systems or devices

[40] These resources are a good starting point to help practitioners understand their roles and responsibilities: NIST SP 800-39, *Managing Information Security Risk-Organization, Mission, and Information System View* (https://csrc.nist.gov/pubs/sp/800/39/final); OMB Circular A-123: *Management's Responsibility for Enterprise Risk Management and Internal Control* (https://obamawhitehouse.archives.gov/omb/circulars_a123_rev); OMB Circular A-130: *Managing Information as a Strategic Resource* (https://www.federalregister.gov/documents/2016/07/28/2016-17872/revision-of-omb-circular-no-a-130-managing-information-as-a-strategic-resource).

Risks **FROM** the data usage *Risks based on how an agency uses or processes data*	• **Data governance risks:** Insufficient or inadequately enforced data governance may lead to non-compliance or mismanagement of data • **Data accountability and stewardship:** Improper data management or accountability, such as a lack of clear data ownership or failing to oversee proper data usage • **Unintended data usage:** Data being used for purposes other than its intended or authorized use, such as using data to instruct AI to manipulate people or systems through deepfakes • **Data aggregation risks:** Unintended consequences from combining multiple datasets, such as privacy violations or misinterpretation of data • **Processing errors:** Mistakes made during data processing, analysis, or transformation, leading to incorrect outputs or conclusions • **Algorithmic bias:** Flawed data can lead to automated data processing systems (e.g., AI) producing biased or unfair results • **Misinterpretation of data:** Incorrect conclusions are drawn from data analysis, leading to improper decision-making • **Privacy risks:** Exposure of sensitive information, such as CUI and PII
Risks **IN** the data *Risks based on the qualities of the data itself*	• **Data entry errors:** Incorrect or inconsistent data input leading to inaccurate or incomplete information • **Outdated data:** Use of obsolete or stale data that no longer accurately reflects the current situation, which may lead to improper decision-making • **Data timeliness risks:** Delayed or slow processing times that result in decisions being made without timely, comprehensive information • **Relevance risks:** Accumulation of unnecessary data or an over-reliance on historical data may distract from key insights • **Unverified sources:** Using data from untrustworthy or unverified sources that lead to false conclusions • **Unreliable data collection methods:** Using data collected through inconsistent or faulty methodologies
Risks **FROM SHARING** data *Risks related to sharing data with agency business partners and third parties, including suppliers with incidental access*	• **Third-party management:** Inadequate security practices by third parties may lead to violations of data's confidentiality, integrity, or availability • **Data misuse:** Data being used for purposes beyond the agreed upon use, potentially violating privacy or compliance requirements • **Data leakage:** Accidental or intentional exposure of shared data during transmission, processing, or storage • **Loss of control:** Diminished ability to monitor and enforce security measures once data is shared with external entities

?

DID YOU KNOW?

Data security practitioners should consult NIST's **Cybersecurity Framework (CSF)** 2.0[41] or the NIST **Risk Management Framework (RMF)**[42] for guidance on effectively addressing specific risk scenarios. However, practitioners should remain cognizant of additional data-centric risk scenarios that may exist outside of these frameworks.

[41] NIST CSF 2.0, https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf

[42] NIST RMF, https://csrc.nist.gov/projects/risk-management/about-rmf

3.2: WHERE TO BEGIN: BRINGING SECURITY TEAMS AND DATA TEAMS TOGETHER

Data teams and security teams can mutually benefit from a collaborative relationship. Data stewards can partner with their security counterparts to build their understanding of security risks and streamline data risk assessments. An agency's security team will likely have standardized tools and processes for assessing risks and may have already quantified risks for certain agency systems.

Likewise, security teams can benefit from data stewards' understanding of the scope of the data landscape and the totality of assets that must be secured. Data stewards should readily know the business value of data assets, which is crucial for accurate identification of data security risks and implementation of appropriate controls.

You don't know what you don't know, so it's essential to work together and bridge potential knowledge gaps to ensure that the agency is implementing a robust data security infrastructure that can mitigate potential data security risks.

You may be able to leverage existing channels for partnership. For example, you can foster collaboration through your agency's formal enterprise-level risk management program, if applicable.

3.3: RISK MANAGEMENT FROM A DATA SECURITY PERSPECTIVE

The NIST RMF provides a holistic process for managing cybersecurity and privacy risks to systems and organizations. Agencies are required to implement the RMF[43] and can further utilize the RMF's methodology to identify data security risks.

When applying the RMF process for data security risk identification, the primary roles will lie with the agency's security team, such as ZT or risk management practitioners. However, the security team should coordinate across their agency to identify common controls[44] and remain flexible in their approach as the roles of data management practitioners may evolve in a ZT data security model.

!

REMEMBER

While it's essential to focus on achieving data security objectives, it's equally important to consider those who will be impacted by your actions. Ensure you communicate with impacted users across your agency and allow sufficient time for feedback to uncover challenges and explore solutions.

See *Appendix D: Roles in Data Security Risk Management* in the companion document for a representative list of roles and responsibility guidance relevant to data security risk management.

[43] Federal agencies are required to implement the NIST RMF, per OMB Circular A-130 (and other guidance from OMB) in fulfillment of FISMA requirements for managing security and privacy risks.

[44] NIST RMF Prepare Task P-5 states that common controls can be identified at different levels of an organization, such as departmental, bureau, subcomponent, individual system, or program area.

3.4: THIRD-PARTY RISKS

Agencies rely on both inter/intra-agency partnerships, as well as commercial organizations, to leverage expertise and innovation to effectively execute priorities and missions. These relationships can result in globally distributed and interconnected network ecosystems spanning **information, communications, and operational technology (ICT/OT**). Third-party relationships are often long-term, trust-based, contractual, and integrated with an agency's daily workflows.

> This guide refers to all external entities as "third party" regardless of agency or commercial status, or relationship within the supply chain.

From a data-centric perspective, sharing data with third parties — whether intentionally or via incidental access — introduces additional risks to the agency with the potential to harm stakeholders, operations, reputation, and resilience. Therefore, agencies must consider the risk and appropriate controls throughout the lifecycle of the relationship with the third party. Learn about example use cases of third-party risk in Table 6.

TABLE 6: Example Use Cases of Third-Party Risk

Agency allows non-government third parties to view and use their data. In this use case, agencies retain full control over the data while allowing third parties to view and use the data.	**Agency contracts non-government third parties to manage their data.** In this use case, third parties act as the custodian for the agency's data assets. The agency may lose some or all control over the data after sharing it.
Agency sends data to non-government third parties for processing. In this use case, the agency loses control of its data after sharing it and relies on the third party to maintain appropriate controls and protections.	**Agency shares data with a variety of third parties, such as: a) other Federal agencies; b) U.S. state, local, tribal, and territorial governments (SLTT); c) foreign governments; etc.** In this use case, the agency loses control of its data after sharing it and relies on third parties to maintain appropriate controls and protections.

Common risks arising from third-party relationships:

- **Non-compliance with data privacy regulations:** Third parties may fail to comply with data privacy regulations, which may subject the agency to a fine or legal dispute.
- **Data residency and sovereignty:** Third parties may store or process data in locations that do not align with the agency's data residency or sovereignty requirements. This can raise concerns about legal jurisdiction, compliance, and control over the data.
- **Business continuity and disaster recovery:** If the third party doesn't have robust business continuity and disaster recovery plans, it can pose a risk to the agency's data availability and recovery in the event of a disruption (e.g., natural disaster, ransomware attack).

3.5: DATA SECURITY THROUGH THE PRIVACY LENS

In today's rapidly evolving digital landscape, **the preservation of individual privacy relies on the protection of sensitive data**. The Federal government's commitment to safeguarding this data is not only a legal obligation as outlined in *Chapter 2*, but a fundamental ethical responsibility, as exemplified in the **Fair Information Practice Principles (FIPPs)**.[45]

Privacy is a cornerstone of a comprehensive data security framework, ensuring that the collection, processing, storage, and dissemination of PII aligns with stringent privacy regulations and ethical considerations. This section will articulate the concerted efforts to navigate the intricate landscape of regulatory requirements, industry standards, and best practices pertaining to data privacy, and aims to guide agencies in the following actions:

- **Review the vast landscape of privacy frameworks requirements.**
    - Ensure compliance with pertinent laws, regulations, and standards governing data privacy.
- **Identify and mitigate privacy risks.**
    - Conduct thorough assessments to identify potential privacy risks and assess their impact.
    - Implement robust mitigation strategies to safeguard sensitive data.
- **Promote "privacy by design."**
    - Integrate privacy considerations into the fabric of your processes, systems, and initiatives.
    - Adhere to the principle of "privacy by design" to embed privacy measures from inception.
- **Collaborate with key stakeholders.**
    - Engage stakeholders, both internal and external, to foster a culture of privacy awareness, education, and collaboration.
    - Recognize that privacy is a collective responsibility.

[45] The FIPPs are a collection of widely accepted principles that agencies use when evaluating information systems, processes, programs, and activities that affect individual privacy. While the FIPPs are not requirements, they are principles that should be applied by each agency according to the agency's particular mission and privacy program requirements. For more information, visit https://www.fpc.gov/resources/fipps/.

3.5.1: INTEGRATING PRIVACY STANDARDS

Privacy standards play a crucial role in shaping data security practices within Federal agencies. They provide a set of guidelines that organizations must adhere to in order to meet legal requirements, ensure data privacy, and mitigate risks.

Practitioners should start by identifying and understanding the specific privacy laws, regulations, and standards that apply to their agency. Agencies may be subject to different laws and regulations based on the type of data they handle (e.g., health, financial, etc.). To ensure compliance, practitioners should know what each regulation authority requires in terms of data protection, privacy safeguards, access controls, reporting, and more.

Conducting a thorough review to understand what each privacy regulation requires can be time-intensive. It may be more efficient for data security practitioners to build strong partnerships and consult with their agency's privacy team in implementing ZT data security.

NIST offers various frameworks and publications that provide guidelines and best practices for cybersecurity, risk management, and data protection. For instance:

- The **NIST Privacy Framework**[46] is complementary to the NIST CSF and assists organizations in managing privacy risks by providing a structured approach to privacy management.
- **NIST SP 800-122, *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)***[47] provides guidance to for protecting the confidentiality of PII in information systems.
- **NIST SP 800-53 Rev. 5, *Security and Privacy Controls for Information Systems and Organizations***[48] provides a catalog of security and privacy controls for Federal information systems and organizations.

Since these controls integrate with privacy, practitioners may find it useful to consult the Federal Privacy Council's security and privacy collaboration index and work with their agency's privacy team when deciding which controls to choose (https://www.fpc.gov/assets/pdf/Collaboration%20Index%20for%20Security%20and%20Privacy%20Controls%20FINAL.pdf).

46 NIST Privacy Framework, https://www.nist.gov/privacy-framework

47 NIST SP 800-122, *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)*, https://csrc.nist.gov/pubs/sp/800/122/final

48 NIST SP 800-53 Rev. 5, *Security and Privacy Controls for Information Systems and Organizations*, https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final

3.5.2: PRIVACY PRESERVATION TECHNIQUES AND TOOLS

The following techniques and tools aim to protect sensitive information, prevent inadvertent disclosure of PII, preserve privacy, minimize the risk of data breaches, and align with the core tenants of ZT.

REMEMBER

It's important to consider appropriate scale — many third-party breaches occur as a result of an organization taking on more data than it can reasonably protect, which has further effects on the third party and the agency in data breach response. Anonymization and pseudonymization are generally used to reduce re-identification risk (i.e., the risk that PII will be revealed through the mosaic effect when combining datasets). Practitioners should note that anonymization does not alleviate all privacy concerns, and it's possible to over-collect anonymized information.

1. **Data minimization and anonymization:** Limit the collection, storage, disclosure, and processing of personal data to only what is necessary for the intended purpose. Anonymize or pseudonymize data whenever possible.
    - **Tools/techniques:** Anonymization tools, tokenization, data masking, k-anonymity, and differential privacy
2. **Encryption:** Use robust encryption methods to protect data at rest, in transit, and in use to prevent unauthorized access even if the data is compromised.
    - **Tools/techniques:** Strong encryption algorithms (e.g., **Advanced Encryption Standard [AES]**, **Rivest-Shamir-Adleman [RSA]**), encryption key management systems, secure communication protocols (e.g., **Transport Layer Security [TLS]**, **Secure Sockets Layer [SSL]**), and secure communications solutions
3. **Data masking and redaction:** Conceal specific portions of sensitive information to prevent unauthorized disclosure while maintaining usability for authorized users.
    - **Tools/techniques:** Redaction tools, dynamic data masking, and techniques to mask or hide sensitive information in documents or databases
4. **Privacy by design and default:** Integrate privacy considerations into the design and architecture of systems and processes by default, ensuring that privacy is a fundamental component of any new initiative.
    - **Tools/techniques:** Design systems with controls in place to protect the data before the system enters into operation, incorporate privacy-enhancing technologies, and adopt privacy-preserving architecture

5. **Privacy-enhancing technologies (PETs):** Leverage technologies explicitly designed to enhance privacy by safeguarding data and minimizing the risks associated with its processing and transmission.
    - **Tools/techniques:** Secure multiparty computation (SMPC), homomorphic encryption, and private information retrieval (PIR)
6. **Compliance automation:** Automate compliance checks and controls to ensure ongoing adherence to privacy regulations and standards.
    - **Tools/techniques:** Compliance and privacy management software, automated risk assessment tools, and policy enforcement mechanisms
7. **Collaboration and standardization:** Collaborate with industry and agency peers and standardization bodies to adopt common privacy-preserving techniques and tools that adhere to recognized standards.
    - **Tools/techniques:** Participation in industry forums, adherence to recognized standards (e.g., those from NIST, International Organization for Standardization [ISO]), and adoption of common best practices

Other approaches, such as computational isolation and confidential computing, can further support preservation of privacy, as outlined in Table 7.

TABLE 7: Computational Isolation and Confidential Computing

Computational Isolation (Clean Data Rooms)	Confidential Computing
Clean data rooms, also known as clean rooms or secure data rooms, can play a crucial role in ensuring security especially within the context of sensitive data handling in Federal agencies. These controlled environments — which can be physical or virtual — provide a secure space for managing highly sensitive information while minimizing the risk of unauthorized access or data breaches. Benefits include: • Secure data handling. • Confidential collaboration. • Regulatory compliance. • Data protection. • Risk mitigation. • Temporary data access. • Comprehensive security measures.	Confidential computing helps keep data safe and private by allowing sensitive information to be processed or analyzed while still encrypted. This means the data stays protected from unauthorized access, even from the system that's processing it. There are emerging technologies that could automatically secure data during processing, ensuring confidentiality, integrity, and privacy. Benefits include: • Protection of sensitive data. • Privacy preservation. • SMPC. • Homomorphic encryption. • Isolation and trusted execution environments (TEEs). • Confidential containers and secure enclaves. • Enhanced compliance and regulatory adherence. • Securing AI models. • Secure cloud and hybrid environments.

3.5.3: PRESERVING PRIVACY WHILE ENGAGING WITH THIRD PARTIES

Practitioners must establish robust protocols, guidelines, and contractual obligations to securely engage with third parties, such as other Federal agencies, U.S. SLTT governments, and foreign governments, and ensure that sensitive data remains protected and compliant throughout the entire supply chain. This may include:

1. **Vendor risk assessments.**
    - **Due diligence:** Conduct thorough assessments to evaluate the security and privacy practices of third parties, especially regarding how they handle sensitive data. Practitioners should also consider whether the third party has the capacity to handle and protect the quantity and sensitivity of data it would receive.
    - **Include privacy criteria:** Integrate criteria related to confidential computing, clean data rooms, encryption, and other privacy preservation techniques into the third-party evaluation process.
2. **Contractual obligations and agreements.**
    - **Data protection clauses:** Include specific clauses in contracts that outline the requirements for protecting sensitive data, specifying the use of confidential computing, clean rooms, encryption standards, and other privacy-preserving techniques.
    - **Compliance adherence:** Ensure third parties commit to complying with relevant U.S. regulations and standards that are applicable to the data they handle.
3. **Security and privacy standards.**
    - **Third-party guidelines:** Provide clear guidelines and standards to third parties regarding the secure handling, processing, and storage of sensitive data.
    - **Training and awareness:** Conduct training sessions or provide resources to educate third parties on best practices for safeguarding sensitive information.
4. **Regular audits and monitoring.**
    - **Continuous evaluation:** Implement regular audits and monitoring mechanisms to ensure that third parties adhere to the agreed-upon security and privacy standards.
    - **Access controls:** Monitor third-party access to sensitive data and ensure strict controls are in place to the extent practicable.
5. **Incident response and reporting.**
    - **Reporting obligations:** Establish clear reporting procedures for any security incidents or breaches and require immediate notification from third parties in case of any data breaches.
    - **Incident response:** Plan a coordinated, timely response in case of security or privacy incidents. Refer to OMB M-17-12, *Preparing for and Responding to a Breach of Personally Identifiable Information* for specific contractor responsibilities.[49]

[49] OMB M-17-12: *Preparing for and Responding to a Breach of Personally Identifiable Information*, https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2017/m-17-12_0.pdf

6. **Legal and compliance oversight.**

   - **Legal review:** Involve legal and compliance experts to ensure that contracts and agreements with third parties adequately address data security, privacy, and compliance requirements.

3.5.4: PRIVACY IMPACT ASSESSMENT

Conducting a **privacy impact assessment (PIA)** involves a systematic evaluation of how a project, system, or initiative might affect the privacy of individuals regarding the collection, use, retention, and disclosure of their PII. It's important to note that a PIA is *both* a process and a document.

For more information about PIAs, refer to the E-Government Act of 2002[50] and OMB M-03-22, *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*.[51]

3.6: DATA SECURITY CONTROLS: IDENTITY, CREDENTIAL, AND ACCESS MANAGEMENT

The philosophy of ZT extends to ICAM, which plays a pivotal role in ensuring the confidentiality, integrity, and availability of data within Federal systems. ICAM encompasses a comprehensive set of policies, procedures, and technologies governing the management of user identities, their access privileges, and the associated activities within an organization's network.

Throughout this section, we will delve into the importance of strong authentication and identity verification, the principle of authorizing access based on least privilege, and the need for continuous monitoring. By understanding and implementing these ICAM principles and controls, agencies can establish a robust and secure environment that minimizes the risk of unauthorized data access, mitigates insider threats, and enhances their overall data security posture.

ICAM practices and technology controls are a primary means of enforcing data security policies for data access and enable the principle of least privilege, both at an information system level and a granular data level.

3.6.1: ACCESS CONTROL MECHANISMS

Access control mechanisms[52] are essential to ICAM with various means of implementation, such as Context-based Access Control (CBAC), Role-based Access Control (RBAC), and Attribute-based Access Control (ABAC). Table 8 explains these access control mechanisms, their pros and cons, and provides example scenarios for choosing the appropriate control. While practitioners should implement controls that best match the data that they are trying

[50] The E-Government Act of 2002, https://www.congress.gov/bill/107th-congress/house-bill/2458

[51] OMB M-03-22: *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*, https://obamawhitehouse.archives.gov/omb/memoranda_m03-22/

[52] Secure Cloud Business Applications: Hybrid Identity Solutions, https://www.cisa.gov/sites/default/files/2023-03/csso-scuba-guidance_document-hybrid_identity_solutions_architecture-2023.03.22-final.pdf

to protect, CISA has observed that access control decisions are best made using CBAC, which considers the context in which the access request is being made. CBAC combines features of RBAC and ABAC to apply dynamic access policies using device-level signals as cues.

TABLE 8: Access Control Mechanisms

Access Control Mechanism	Description and Guidance
Role-based Access Control (RBAC)	Defines access privileges based on job roles and responsibilities. • When using RBAC, data owners and operators can use ISACA's step-by-step guide for establishing appropriate **separation of duties (SoD)**.[53] • Some products can help identify where SoD should be applied by using "role mining" techniques. Data owners and operators should ask if their organizations' ICAM capabilities offer support for role mining if they are having difficulty understanding how to create appropriate roles.
Attribute-based Access Control (ABAC)	Generally used for more granular and dynamic access control based on user attributes and contextual information. • When using ABAC, you can consider attributes like user location, time of access and other factors to determine access privileges so long as the attributes are available at the time the decision is made. • The **National Security Agency (NSA)** and CISA recommend "utilizing a data tagging system or solution, where data is conditionally accessed via granular ABAC policies to protect data. It is also important to separate accounts that grant access to resources from those that manage them daily."[54]
Context-based Access Control (CBAC)	Unlike traditional access control methods, CBAC continuously evaluates factors such as user identity, location, device health, and behavior to make precise access decisions. This method enhances security by ensuring that access is granted only under secure and appropriate conditions. • Ensure that CBAC is integrated into your broader ZTA. This means verifying every access request explicitly, using least privilege principles, and assuming breach to minimize potential damage.

[53] A Step-by-Step SoD Implementation Guide, https://www.isaca.org/resources/isaca-journal/issues/2022/volume-5/a-step-by-step-sod-implementation-guide

[54] Secure Data in the Cloud, https://media.defense.gov/2024/Mar/07/2003407862/-1/-1/0/CSI-CloudTop10-Secure-Data.PDF

3.6.2: ESSENTIAL IDENTITY, CREDENTIAL, AND ACCESS MANAGEMENT PRACTICES FOR PROTECTING DATA

While there are many factors that practitioners should evaluate when determining how to properly and appropriately protect data, this guide considers the ICAM practices outlined in Table 9 to be essential.

TABLE 9: Essential Identity, Credential, and Access Management Practices for Protecting Data

ICAM Practice	Description
Identity validation and verification	Involves the processes and technologies used to confirm that an individual's claimed identity is genuine and accurate.
Principle of least privilege	Mandates users, applications, and systems be granted the minimum level of access necessary to perform their functions. By limiting access rights, it reduces the risk of unauthorized access and potential security breaches, ensuring that users can only access the information and resources essential for their legitimate purposes.
User provisioning and lifecycle management	Involves the processes and technologies used to manage user identities and their access to resources throughout their entire lifecycle within an organization. This includes onboarding new users, managing changes in user roles, and deprovisioning users when they leave the organization.
Continuous monitoring	An ongoing, systematic process that utilizes automated tools to constantly observe and analyze an organization's IT systems and networks. The primary goal is to detect security threats, performance issues, and compliance problems in real time, allowing for prompt identification and resolution of potential risks.
Authentication	The process of verifying the identity of users, devices, or systems before granting them access to resources. It ensures that only authorized entities can access sensitive information and systems, thereby protecting against unauthorized access and potential security breaches.
Access control mechanisms	The methods and processes used to regulate who or what can view, use, or access resources within a system. These mechanisms are essential for ensuring the security and integrity of data by preventing unauthorized access and managing permissions based on user roles and contexts.
Federation and single sign-on	Authentication mechanisms are designed to simplify and secure access to multiple applications and systems. Federation allows users to access resources across different domains or organizations using a single set of credentials, while **single sign-on (SSO)** enables users to log in once and gain access to multiple applications within the same domain without needing to re-authenticate.

3.7 DATA SECURITY MONITORING AND CONTROLS

This section provides guidance on enhancing data security and implementing rigorous monitoring for data — whether it's at rest or in transit — within a ZT framework. We will delve deeper into the realm of data-centric security monitoring, logging, and alerting. These are crucial components in identifying and responding to anomalous or suspicious activities that could potentially compromise data security.

3.7.1: USE DATA-CENTRIC SECURITY CONTROLS TO SECURE DATA AT EVERY LEVEL, IN EVERY LOCATION

Cybersecurity technologies from firewalls to endpoint security tools are used to help protect data, but data-centric security controls consist of a specific domain of technologies that bring the controls closer to the data itself. Common examples of data-centric security controls include encryption, rights management, and **data loss prevention (DLP)**. Data access controls, such as use of RBAC, CBAC, and ABAC are used in conjunction with data-centric security controls to ensure appropriate levels of access and data use. The controls an agency selects to use should be chosen to mitigate the data risks identified, protect the data while enabling appropriate use based on its sensitivity level, and achieve specific objectives.

3.7.2: DEFINE POLICIES AND SELECT THE APPROPRIATE CONTROLS

Data security controls are used to enforce policies. Practitioners should define their policies based on data sensitivity, identifying who can access the data, under what conditions, and what they can do with it.

Access enforcement can limit a threat actor from interacting directly with a protected resource or data, while segment enforcement limits the ability for a threat actor to create further loss to the organization after breaching access and perimeter controls. Practitioners should expect to use a combination of various data-centric security controls and data access controls to address key objectives as outlined in Table 10.

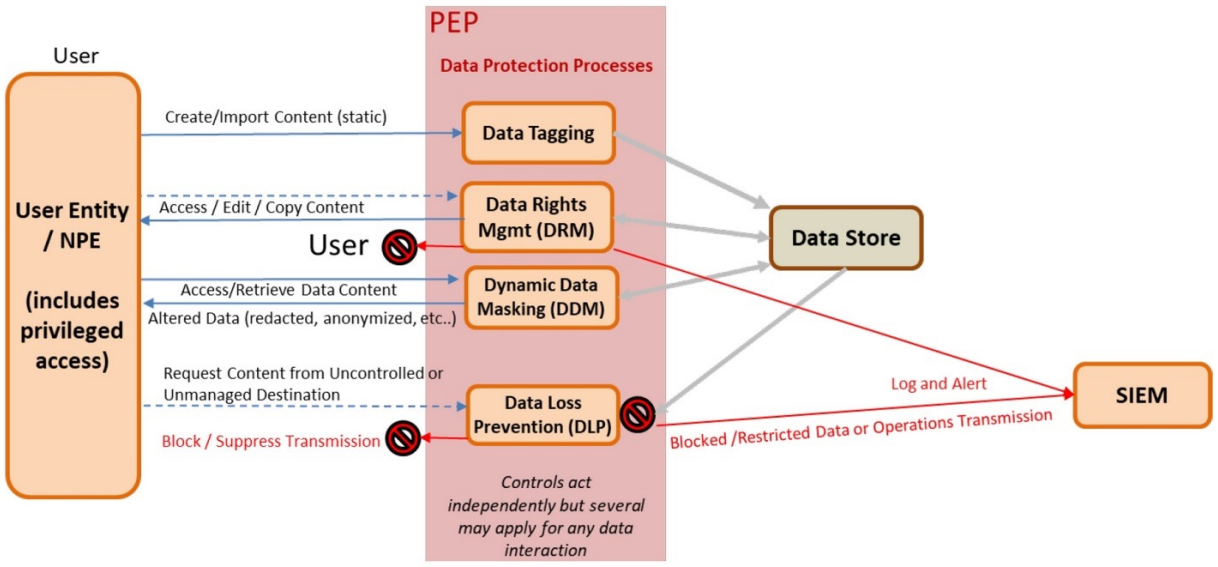
TABLE 10: Data-Centric Control Type Objectives

Control Type Objective	Example Technology Control
Reduce unauthorized access	Data access controls (e.g., RBAC, CBAC, ABAC), data access governance, rights management, data masking, data redaction, confidential computing
Reduce loss event frequency	DLP, data detection and response
Reduce loss magnitude	Encryption, tokenization, data masking, data redaction
Provide visibility into data and user activity	Database activity monitoring, DLP, data detection and response, **user and entity behavior analytics (UEBA)**
Preserve privacy	Privacy-preserving technologies, clean data rooms, confidential computing

For descriptions of example data-centric security controls, see *Appendix C: Security Monitoring and Controls* in the companion document.

Aligning technical policies with data access needs enables **policy enforcement points (PEPs)** to provide robust, data-centric protection. In implementing ZT, agencies would use a policy engine to deploy the policy and enforce the policy through a PEP. Access requests would be evaluated against the policy and adjudicated in real time. This approach ensures that sensitive data is adequately protected, supporting regulatory compliance, and reducing the risk of data breaches. See Figure 4 for an agency example of this alignment and use of data-centric security controls.

FIGURE 4: Data-Centric Security Protections



Source: Department of Defense Zero Trust Reference Architecture, https://dodcio.defense.gov/Portals/0/Documents/Library/(U)ZT_RA_v2.0(U)_Sep22.pdf

3.7.3: CONTINUOUSLY MONITOR FOR SECURITY CONTROL EFFECTIVENESS

Continuous monitoring capabilities across several key areas will help agencies stay on top of their changing data environment, ensure ongoing alignment between policies and controls, and respond quickly to suspicious activity, as outlined in Table 11.

TABLE 11: Continuous Monitoring Capabilities

Capability	How To Enable and Operationalize
Automated data inventory scanning. This ensures that the data inventory remains up to date. This capability typically includes tools and scripts to scan an organization's network, file systems, databases, and cloud repositories to discover data assets and perform initial identification.	• Data inventory scanning is often a feature within data security tools, referred to as data discovery functionality. Data repositories may also have this capability. It is also available as a standalone technology. • When new data assets are discovered or when existing assets change, use automated alerts to notify data stewards or administrators. This enables timely action to properly inventory and secure the data.
Data monitoring. This enables immediate alerts for potential security incidents. It includes real-time monitoring of data access, data activity and modifications, and user behavior. It also includes anomaly detection to identify suspicious activity.	• DLP capabilities are a common approach for data monitoring to alert on policy violations related to data movement. • Identify security analysts responsible for triaging and investigating alerts generated from monitoring tools. • With the use of monitoring technologies, ensure proper considerations and communications related to employee privacy. Develop procedures for an appropriate response to alerts, particularly as it relates to addressing insider risks.

3.7.4: LOG, AUDIT, AND ALERT TO MAINTAIN SECURITY AND ENABLE INVESTIGATIONS

Data-centric logging, auditing, and alerting are crucial for ZT, especially from the PEP. The PEP is responsible for enforcing the organization's security policies, and its logs are vital for maintaining security. Important logs from the PEP include:

1. **Access logs:** These logs record who accessed what data, when, and from where.
2. **Policy decision logs:** These logs track the decisions made by the PEP, such as allowing or denying access based on the policies.
3. **Change logs:** These logs track any changes made to the policies or system configurations.

Auditing involves regularly reviewing these logs to identify any anomalies or suspicious activities. Automated tools can help in this process by flagging activities that deviate from established norms or violate security policies. Audit logs track user actions and system changes to ensure accountability and traceability. They provide a chronological record of activities, crucial for audits and compliance checks. They may be used for evidence purposes in compliance, security, and computer forensic investigations.

Alerting is the process of notifying relevant personnel when such anomalies are detected. This allows for immediate action to mitigate potential threats.

The Role of the Security Operations Center in Monitoring and Response

Ideally, agencies will have a high-functioning **Security Operations Center (SOC)** to play a pivotal role in data security efforts. SOCs continuously monitor network traffic, user behavior, and system activities for any signs of anomalous or suspicious activity. They use advanced analytics and threat intelligence to identify potential threats that may otherwise go unnoticed. In the absence of an SOC or existence of an SOC with limited staffing, agencies may rely on a **managed security services provider (MSSP)** for this function. **Managed detection and response (MDR)** providers are an option for organizations to obtain 24/7 monitoring of their security events and augment existing SOC analysts. Alternatively, agencies can visit https://www.cisa.gov/resources-tools/services/security-operations-center-soc-optimization-advisory-service to learn more about their SOC Optimization Advisory Service.

DID YOU KNOW

Tabletop exercises are crucial for enhancing data security by simulating scenarios that test an agency's response to data breaches and cyber threats. These exercises allow teams to analyze how different security measures intersect and impact overall data protection strategies. By involving diverse stakeholders, agencies can identify vulnerabilities, assess the effectiveness of their security protocols, and develop comprehensive response plans. This proactive approach ensures that data security measures are robust, inclusive, and capable of addressing the complex challenges posed by modern cyber threats.

3.8 PRACTICAL EXAMPLE: STEPS FOR POLICY ENFORCEMENT CONTROLS

To effectively monitor data usage and detect any suspicious activities, technical practitioners should implement robust security measures. Here's a practical example to apply these principles:

Step 1: Prioritize known data types. Begin by focusing on well-defined data categories, as defined in Chapter 2, such as files containing PII like **social security numbers (SSNs)**, personnel forms (e.g., SF-50), or CUI. These data types are commonly targeted and should be secured first.

Step 2: Identify risks to mitigate. Determine the purpose for the security control and why it is necessary, alongside an understanding of how the data needs to be used or flow. For example, the agency needs to share files containing CUI with third parties, and employees should use the agency-approved, secure file-sharing solution to do this rather than email or public cloud storage services.

Step 3: Use security controls to enforce policies. Implement security controls to prevent data from being transferred from a secure environment to a less secure one. For instance, if data is being sent from an internal network to an external one (e.g., via email or to cloud storage services like Dropbox), it should trigger a security protocol. In this example, use DLP tools to automate detection and response to this policy violation.

Step 4: Adopt an iterative approach. Start with basic actions such as logging any data movement attempts. Gradually escalate the response to include user notifications, alerts to an SOC, data transfer quarantines, and — eventually — blocking unauthorized transfers.

See *Appendix C: Security Monitoring and Controls* in the companion document for an example implementation plan.

CHAPTER 4: MANAGE THE DATA

Placeholder Notice – We've included a placeholder for this chapter as we continue to develop and refine content. The goal of this chapter will be to ensure that data security practices are aligned with and embedded in DLM. Our priority is to share insights at the speed of need, and we hope that practitioners can leverage the previous chapters to begin the process of defining and securing their data.

CHAPTER 5: CONCLUSION

5.1: RECAP

This guide outlines an approach to implementing ZT data security within the Federal government. We began by establishing the vision and principles of ZT, emphasizing the need to view data as the new perimeter. We focused on how to define data using various types, classifications, and categories before expanding on how to secure data with the appropriate security controls.

5.2: BUILDING THE FUTURE OF DATA SECURITY

As we look to the future of data security, it's clear that we must embrace the following to be successful:

- **Cross-functional collaboration and effective communication:** Cybersecurity is a collective effort that goes beyond agency IT and cybersecurity teams — it requires collaboration across all roles at all levels. Effective communication between teams helps ensure alignment and promote interconnectedness, fostering a culture of security where everyone contributes to the collective effort of protecting the organization's digital assets. This also enables continuous improvement as different roles can learn from each other, share best practices, provide feedback on existing policies, and collaborate on new security initiatives.
- **A cooperative relationship between data and security teams:** Fostering a CoP between data stewards and security practitioners is paramount. A persistent, collaborative environment encourages the exchange of knowledge, promotes the development of innovative solutions, fosters the development of a culture of security, and strengthens an agency's overall security posture. CoPs play a crucial role in this process, serving as a platform for continuous learning and improvement, thereby ensuring that the principles of ZT are not statically implemented, but evolve with the changing cybersecurity landscape.
- **Continuous learning and education:** Since EO 14028 and OMB M-22-09, there has been a significant increase in ZT training available to multiple audiences. Agencies and departments should invest in training and certification programs to develop cross-functional teams with common lexicons and language. This applies to all levels of the Federal workforce, particularly senior leadership, to enable them to make informed decisions that balance the benefits and risks of data-driven initiatives. Additionally, new resources and frameworks are being developed that offer innovative ways to protect and analyze data. Practitioners should endeavor to stay informed about the latest developments. By fostering a culture of continuous learning and awareness, agencies

can equip their personnel with the necessary skills to navigate and counter evolving cyber threats.

- **Intersectional analysis and assessments:** Further consideration should be provided to the data that departments and agencies already collect, supporting greater fidelity of data and associated attributes.
- **Across-the-board buy-in:** It's crucial to gain support from policy officials and senior leaders, Congress, and other stakeholders to adequately prioritize funding and resources for ZT data security efforts. Stakeholders should engage in multi-year planning to ensure their agency can keep up with the pace of cybersecurity over time.
- **Adaptability:** Practitioners need to find the right implementation strategy that meets the unique needs of their agency. Implementation strategies should be approached from both a strategic and a tactical perspective. There is no one-size-fits-all solution or tactical level implementation strategy.

In today's interconnected world, all devices will — or already do — seamlessly communicate with each other, **expanding the threat surface of an already vast data ecosystem.** Therefore, the implementation of ZT principles is paramount for the Federal government to safeguard its data assets in an increasingly complex and contested cyber environment. **By adhering to the core tenets of ZT — never trust, always verify, and assume breach — agencies can ensure that their data is categorized and handled with the utmost precision and care.**

Agencies must ultimately embrace innovation and collaboration to mature the data security architecture of the Federal enterprise and build resilience against future threats.


federal
CDO
council
CISO
Council
CIO.GOV

Federal Zero Trust Data Security Guide: Appendices

OCTOBER 2024

federal CDO council | CISO Council | CIO.GOV

TABLE OF CONTENTS

TABLES

RESOURCES

CISA PUBLICATIONS

- CISA Zero Trust Maturity Model, https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf
- CISA BOD 18-02: Securing High Value Assets, https://www.cisa.gov/news-events/directives/bod-18-02-securing-high-value-assets

EXECUTIVE ORDERS

- EO 14028 on *Improving the Nation's Cybersecurity*, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/
- EO 13744 on *Making Open and Machine Readable the New Default for Government Information*, https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

LAWS

- FISMA 2014, https://www.congress.gov/bill/113th-congress/senate-bill/2521
- HIPAA 1996, https://www.congress.gov/bill/104th-congress/house-bill/3103/text
- The E-Government Act of 2002, https://www.congress.gov/bill/107th-congress/house-bill/2458
- The Foundations for Evidence-Based Policy Making Act of 2018, https://www.congress.gov/bill/115th-congress/house-bill/4174
- The Paperwork Reduction Act of 1980, https://www.congress.gov/96/statute/STATUTE-94/STATUTE-94-Pg2812.pdf
- The Privacy Act of 1974, https://www.govinfo.gov/content/pkg/USCODE-2018-title5/pdf/USCODE-2018-title5-partI-chap5-subchapII-sec552a.pdf

NIST PUBLICATIONS

- FIPS 199, *Standards for Security Categorization of Federal Information and Information Systems*, https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.199.pdf
- NIST Cybersecurity Framework (CSF) 2.0, https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf

- NIST Privacy Framework, https://www.nist.gov/privacy-framework
- NIST Risk Management Framework (RMF), https://csrc.nist.gov/projects/risk-management/about-rmf
- NIST SP 800-39, *Managing Information Security Risk: Organization, Mission, and Information System View*, https://csrc.nist.gov/pubs/sp/800/39/final
- NIST SP 800-53 Rev. 5, *Security and Privacy Controls for Information Systems and Organizations*, https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final
- NIST SP 800-60, *Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories*, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v2r1.pdf
- NIST SP 800-63, *Digital Identity Guidelines*, https://pages.nist.gov/800-63-3/
- NIST SP 800-122, *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)*, https://csrc.nist.gov/pubs/sp/800/122/final

OMB CIRCULARS

- OMB Circular A-123, *Management's Responsibility for Enterprise Risk Management and Internal Control*, https://obamawhitehouse.archives.gov/omb/circulars_a123_rev
- OMB Circular A-130, *Managing Information as a Strategic Resource*, https://www.federalregister.gov/documents/2016/07/28/2016-17872/revision-of-omb-circular-no-a-130-managing-information-as-a-strategic-resource

OMB GUIDANCE

- OMB M-03-22: *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*, https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2003/m03_22.pdf
- OMB M-16-17: *OMB Circular No. A-123, Management's Responsibility for Enterprise Risk Management and Internal Control*, https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2016/m-16-17.pdf
- OMB M-17-12: *Preparing for and Responding to a Breach of Personally Identifiable Information*, https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2017/m-17-12_0.pdf
- OMB M-18-16: *Appendix A to OMB Circular No. A-123, Management of Reporting and Data Integrity Risk*, https://www.whitehouse.gov/wp-content/uploads/2018/06/M-18-16.pdf
- OMB M-19-03: *Strengthening the Cybersecurity of Federal Agencies by enhancing the High Value Asset Program*, https://www.whitehouse.gov/wp-content/uploads/2018/12/M-19-03.pdf
- OMB M-19-18: *Federal Data Strategy — A Framework for Consistency*, https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf

- OMB M-21-27: *Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans*, https://www.whitehouse.gov/wp-content/uploads/2021/06/M-21-27.pdf
- OMB M-22-09: *Moving the U.S. Government Towards Zero Trust Cybersecurity Principles*, https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf
- OMB M-23-04: *Establishment of Standard Application Process Requirements on Recognized Statistical Agencies and Units*, https://www.whitehouse.gov/wp-content/uploads/2022/12/M-23-04.pdf
- OMB M-24-14: *Administration Cybersecurity Priorities for the FY 2026 Budget*, https://www.whitehouse.gov/wp-content/uploads/2024/07/FY26-Cybersecurity-Priorities-Memo_Signed.pdf

OTHER

- A Framework for Data Quality, https://nces.ed.gov/FCSM/pdf/FCSM.20.04_A_Framework_for_Data_Quality.pdf
- Business Reference Model of Federal Enterprise Architecture, https://obamawhitehouse.archives.gov/omb/e-gov/FEA
- CDO Council Data Inventory Report, https://resources.data.gov/assets/documents/CDOC_Data_Inventory_Report_Final.pdf
- DATA Act Information Model Schema (DAIMS), https://resources.data.gov/standards/catalog/daims/
- DCAT-US v1.1, https://resources.data.gov/resources/dcat-us/
- DCAT-US v3.0 Schema, https://github.com/DOI-DO/dcat-us
- Desirable Characteristics of Data Repositories for Federally Funded Research, https://www.whitehouse.gov/wp-content/uploads/2022/05/05-2022-Desirable-Characteristics-of-Data-Repositories.pdf
- Digital Identity Risk Assessment (DIRA), U.S. Agency for International Development, https://www.usaid.gov/digitalstrategy/dira
- General Records Schedules, https://www.archives.gov/records-mgmt/grs
- Guidance for Assigning New Cybersecurity Codes to Positions with Information Technology, Cybersecurity, and Cyber-Related Functions, CHCO Council, https://www.chcoc.gov/content/guidance-assigning-new-cybersecurity-codes-positions-information-technology-cybersecurity
- Guide for Mapping Types of Information and Systems to Security Categories, https://doi.org/10.6028/NIST.SP.800-60r2.iwd
- Joint Knowledge Online Login; Search for DoD course “US005” — DAU ZT for Executives, https://jkodirect.jten.mil/Atlas2/page/login/Login.jsf

- National Cyber-Informed Engineering Strategy, U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response, https://www.energy.gov/sites/default/files/2022-06/FINAL%20DOE%20National%20CIE%20Strategy%20-%20June%202022_0.pdf
- National Cyber Workforce and Education Strategy, Office of the National Cyber Director, https://www.whitehouse.gov/wp-content/uploads/2023/07/NCWES-2023.07.31.pdf
- National Cybersecurity Strategy, https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf
- NIEMOpen, https://www.niem.gov/

GLOSSARY

Data and security practitioners must understand each other's nomenclature to effectively safeguard their agencies' data and enable appropriate use. Table 1 outlines terms that can have slight variations in definition depending on their interpretation from a security or data management perspective.

TABLE 1: Glossary

Term	Definition from Security Perspective	Definition from Data Perspective
Data discovery	Identifying where sensitive data is located, such as endpoint and cloud applications.	Finding and cataloging data sources, such as databases, data warehouses, and data lakes. You may see this referred to as data ingestion.
Data categorization	This is sometimes used interchangeably with data classification. It may be used to describe data type, such as "invoice," "memo," or "intake form."	Identifying the categories of the data assets.
Data classification	Identifying what data is sensitive and tagging/labeling that data, then determining the risk to the organization if the data is exposed by sensitivity level. The purpose is to inform the appropriate data controls (including automation), handling practices, and third-party requirements.	Identifying data type.
Data tagging or labeling	This may be used in reference to data classification, with tags/labels, such as "public" vs. "internal" vs. "highly restricted." It could reference tags/labels such as "PII" vs. "Finance data" vs. "HR data." The purpose is to inform the appropriate data controls (including automation) and handling practices.	Tagging/labeling for informing business use purposes, such as use of tags in a data catalog. *See Appendix A for more details.*
Data inventory	Discovery, categorization, classification, and mapping of data flows. You'll often see this term used in the context of privacy compliance.	Making a list of all data assets and gathering the technical metadata. *See Appendix A for more details.*

Data governance	This can be viewed as security and privacy controls and processes, though this is only one aspect of data governance.	An umbrella term and theme covering multiple competencies, including master data management, data quality, data retention, data integrity, data democratization, data literacy, and data compliance and risk. It may also include capabilities like data definitions, policies, quality, stewardship, literacy, regulatory requirements, ethical considerations, risk management, privacy and security, and end-to-end lifecycle management.
Data integrity	Preventing data tampering or manipulation.	Related to data accuracy and quality.
Data protection	Often viewed as data security controls, in addition to related data handling processes/ practices meeting privacy compliance requirements. Focus is primarily on ensuring data confidentiality.	This could be viewed as synonymous with data security.

APPENDIX A: DATA INVENTORY

A.1: A DATA INVENTORY AND DATA CATALOG ARE COMPLEMENTARY YET DISTINCT

While "data inventory" and "data catalog" are often used interchangeably, it's important to understand that they are not the same. In April 2022, the CDO Council released the Enterprise Data Inventories Report[1] which highlights the distinction between these two artifacts:

> "Although the terms 'data catalog' and 'data inventory' are often used synonymously, they mean very different things. A data catalog is the mechanism that helps users discover the data assets that are found in the data inventory. The data catalog contains such information as the organizational ownership of data assets, its meaning to the organization (business metadata), and where and how to access it. However, a data inventory can contain more technical and granular metadata such as the definitions of specific data elements, their format, valid values, and their completeness. While the concepts are distinct, they are complementary."

The report notes that agencies should plan on maintaining *both* a data inventory and a data catalog.

Table 2 provides a summary of the differences between a data inventory and data catalog.

TABLE 2: Data Inventory vs. Data Catalog: Definitions

Aspect	Data Inventory	Data Catalog
Definition	Details the type and location of each data point in an organization	References an organization's datasets in various categories for search and discovery
Scope	Helps map an organization's data, primarily for compliance with regulations	Enables data search and discovery of data assets, with the right context. It also ensure data quality, integrity, and reliability
Users	IT teams use it to find and map all essential data assets	Technical and business users use it to access the right data and extract insights
Key difference	Includes the technical metadata associated with each data asset	Includes all metadata types — technical, business, operational, and social

[1] CDO Council Data Inventory Report, https://resources.data.gov/assets/documents/CDOC_Data_Inventory_Report_Final.pdf

Top benefits	**Transparency and awareness** IT teams know what data their organizations collect, store, and use, including dark data	**A single source of truth** Serves as a central repository for everyone within an organization to find and access the data they need
	Trustworthy data IT teams can trace data origins and verify its authenticity and credibility	**High-quality, timely, and trustworthy data** Automates lineage and propagates policies through lineage; creates automatic data profiles and runs automated quality checks frequently to spot anomalies or inconsistencies in data
	Legal compliance for sensitive data Helps with legal and regulatory compliance by finding and mapping sensitive data	**End-to-end governance and data democratization** Helps with compliance by enabling granular (column-level) access controls, lineage mapping, tag-based access policies, and automated PII data classification
Relationship	A data inventory involves identifying all the data of an organization. It is the first step toward creating a data catalog	Inventorying data is an essential aspect of data catalogs. They're created after identifying the data within an organization's warehouses and lakes

A.2: FEDERAL REQUIREMENTS TO CREATE AN ENTERPRISE DATA INVENTORY

The need to maintain an accurate data inventory and catalog is emphasized in multiple Federal documents and communities. The CDO Council's report, "Enterprise Data Inventories: Agencies face challenges and opportunities to increase the value of data assets when implementing data inventories," provides a comprehensive list of statutory requirements for creating an enterprise data inventory.

*The **Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act)** calls for a systematic rethinking of how the Federal government manages and uses the information it collects, emphasizing strong agency coordination for the strategic use of data. Specifically, Title II of the Evidence Act is the **Open, Public, Electronic and Necessary (OPEN)** Government Data Act (an Act within an Act), which lays out certain agency responsibilities, including the requirement for a comprehensive data inventory.*

A.3: THE FOUNDATION OF AN ENTERPRISE DATA INVENTORY AND ENTERPRISE DATA CATALOG

Oversight/governance authority. Creating an accurate data inventory requires the support and accountability of all stakeholders. A multi-disciplinary, cross-functional agency group that manages the enterprise data inventory to ensure uniform processes increases the likelihood of successful agency adoption.

Identify metadata (i.e., data about data assets). Agencies should use a business-oriented metadata collection approach when conducting data inventory and catalog. The metadata should include data processes, sources, purposes, storage locations, subjects, types, and why, where, and how the data is entering the agency's system. This metadata allows the agency to craft policies, ensure compliance with applicable laws, and understand the risk to, from, and in the data. *See Chapter 3.1 in the Federal Zero Trust Data Security Guide for more details about data risks.*

Periodically review workflows and practices. To prevent the inventory and catalog metadata from becoming stale and losing value, it is essential to conduct regular reviews of workflows and practices. This maintenance is also crucial for ensuring that data security policies are up to date, align with current use cases, and meet changing requirements and needs. A data inventory and catalog are snapshots in time and — since data is constantly moving through its lifecycle – the inventory and catalog metadata will likely change.

A.4: PRIORITIZING THE MANUAL POPULATION OF AN ENTERPRISE DATA INVENTORY OR ENTERPRISE DATA CATALOG

Practitioners should begin with what may already exist inside their agency. If there is no automation tool to create the inventory and catalog, practitioners will need to manually create the data inventory and then move to the data catalog. The agency's privacy team may already have a **Governance, Risk, and Compliance (GRC)** tool populated with a starter set of data catalog information. At a minimum, the GRC tool should identify all agency **high value assets (HVAs)**. Agency HVAs and their associated data assets, typically structured and semi-structured data, should be the first entries in the data inventory and data catalog. Practitioners may also prioritize manually populating the data inventory and data catalog with the FISMA risk level data.

A.5: STEPS TO MANUALLY CREATE AN INITIAL ENTERPRISE DATA INVENTORY OR ENTERPRISE DATA CATALOG WHEN THE AGENCY HAS A GOVERNANCE, RISK, AND COMPLIANCE TOOL

1. Consult the privacy team and a technical data steward with access to the agency's GRC tool to obtain a report of the agency HVAs and the contact information of responsible technical data stewards and technical **subject matter experts (SMEs)**.
2. Collaborate with the technical data stewards and SME(s) to obtain the data inventory for the system and gather metadata.
3. Complete data gathering for all HVAs and put the information in a common tool (e.g., Word, Excel, or CSV file).
4. Repeat the above steps for all FISMA High Impact systems, then Moderate Impact systems, and lastly Low Impact systems.
5. Establish a process for periodic review of the data inventory to ensure that it stays accurate.

If a practitioner's agency doesn't have the budget to purchase a new catalog software, there are other ways to get the process started at low cost by investing processing time to collect the information. These steps are essential so that the agency can categorize and label its data assets and then subsequently apply the appropriate security controls to protect the data assets.

The shortcoming of this manual approach is that the inventory will be limited to known data (typically structured and semi-structured) as identified in the agency's GRC tool. This may leave the inventory incomplete. It also does not provide any coverage for structured and unstructured data that are stored on on-premises agency mass storage devices, cloud locations, etc.

A.6: STEPS TO MANUALLY CREATE AN INITIAL ENTERPRISE DATA INVENTORY OR ENTERPRISE DATA CATALOG WHEN THE AGENCY DOES NOT HAVE A GOVERNANCE, RISK, AND COMPLIANCE TOOL

1. Consult the agency's technical data steward(s) to create a technical metadata[2] inventory. Enter the technical metadata inventory into a common tool (e.g., Word, Excel, or CSV document).
2. Consult the agency's business data steward(s) to create an operational metadata[3] inventory. Enter the operational metadata inventory into a common tool (e.g., Word, Excel, or CSV document).
3. Consult the agency's business data steward(s) to create a business metadata[4] inventory. Enter the business metadata inventory into a common tool (e.g., Word, Excel, or CSV document).
4. Merge the technical, operational, and business metadata inventory files into a single standard file.
5. Establish a process for periodic review of the data inventory to ensure its accuracy.

[2] IT stewards or technical users capture technical metadata. It includes descriptions of existing data assets (e.g., database, data lake, other storage layer or system), data structures (e.g., models, schemas), and data locations (e.g., virtual private cloud [VPC], on-premises). Technical metadata is extracted through data or schema exports from database management systems or other data management platforms like data integration, data quality, or Master Data Management (MDM) tools. Data & Analytics leaders and their teams use existing metadata export functions or manually capture metadata visually from user interfaces or queries to APIs. It includes file formats, media specifications, and source information, shedding light on the inner workings of the data. It also includes identifying systems, applications, databases, sources, targets, security controls, etc.

[3] Operational metadata is extracted from log files from transactional systems or data management tools and differs from technical metadata in that it is created whenever actions are taken on data. For example, if data is moved, copied, transformed, or otherwise updated, this creates operational metadata. Operational metadata is captured in a similar manner to technical data (i.e., via exports, UIs, and APIs) by Data & Analytics leaders and their teams. Log files are a common source, as data movement generates logged transactions and events. Operational metadata commonly references technical metadata as sources and targets of data actions.

[4] Business metadata is captured by business users, analysts, line-of-business owners, and enterprise architects. It starts with business glossaries, where business teams capture standard business terms, processes, and other definitions to be shared and standardized. Business metadata defines not only the business terms but also describes processes in a business context. It refers to the information that provides context and insights into how the data is used within the organization and represents the data's characteristics, origins, and relationships, enabling users to understand and interpret it accurately. Business metadata encompasses various aspects, including data definitions, lineage, classification, and governance policies.

A.7: PRIORITIZING THE INITIAL AUTOMATED DISCOVERY OF AGENCY DATA IN A DATA INVENTORY OR DATA CATALOG

If a practitioner's agency has the budget to purchase a Data & Analytics Governance tool, document the business and technical requirements to ensure alignment with mission-oriented goals and strategic objectives and compatibility with the technology landscape, such as data management tools, GRC tools, and network security boundaries. The requirements should include use cases encompassing critical data management needs, such as research and data analysis, data issue identification, and compliance management. In addition to requirements gathering, practitioners should establish standards, definitions, policies, and processes for entering data assets into the data inventory or catalog tool.

An automated tool allows for additional capabilities that enable the classification of the data against an agency's data taxonomy or data classification schema. This data taxonomy should be entered or built into the data inventory or catalog, and data stewards should be assigned by those defined data subject areas or data concepts. Once those foundational steps are complete, the agency's HVAs and their associated data assets should be entered into the automated data inventory or catalog in alignment with priorities set forth by the agency's requirements and use cases. Similar to the manual process, prioritize the population of the data inventory or catalog by information or data type with the FISMA risk level.

A.8: STEPS TO AUTOMATICALLY CREATE AN ENTERPRISE DATA INVENTORY OR ENTERPRISE DATA CATALOG

1. Run a report with the assistance of the agency's privacy team and technical data steward(s) who have access to the agency's GRC tool or a system inventory tool that includes risk information. This report will identify the agency's HVAs, the associated data stewards and technical SMEs, and the access information for each data source.

2. Establishing a robust technical architecture that depicts the data inventory or catalog components is crucial. This architecture should be designed to address the agency's data sources and outline the methods for moving the inventory from these sources into the catalog tool.

3. When planning and implementing the data inventory or catalog tool, selecting the appropriate development framework is essential. This includes building the front-end application with critical features and capabilities. These features, such as repository, search, data asset linkages, workflows, and metadata extractions, should strongly focus on data security, privacy, integrity, and compliance. This can be achieved through appropriate encryption, masking, and authorization mechanisms.

4. Establish the enterprise business glossary or the agency's data taxonomy (i.e., a collection of business terms organized into a hierarchical structure, typically parent-child relationships, used to classify the types of data and information collected by the agency).

5. Enter any relevant data standards, policies, rules, and other classification assets and information into the data inventory or catalog to establish a relationship between these items and the data inventory from source systems.

6. Working with each system's technical data stewards and SMEs, confirm the existence of the system's data inventory and plan the approach to accessing and obtaining it.

7. Consult the technical data stewards and technical SMEs to determine the appropriate method to extract the metadata for the data inventory. Assess the metadata volume, the storage location format, compatibility with the data inventory or catalog tool and its environment, and other import/export requirements — leverage tool capabilities to auto-discover and auto-populate metadata from sources. While the data gathering may still need curation, this can help reduce the workload upfront. Catalogs must be able to connect to and capture metadata directly from the sources.

8. Export the metadata from the source system to the data inventory or catalog platform in order of HVA systems, FISMA High Impact, FISMA Medium Impact, FISMA Low Impact systems, and other low priority systems per agency requirements and use cases.

9. Ensure the data inventory or catalog platform enables data stewards to review all their metadata periodically to ensure accuracy and timeliness. Determine an appropriate frequency for refreshes or automated updates.

10. Business/technical data stewards and SMEs should create or validate relationships between their data inventory and the agency data taxonomy, data standards, data policies, and other related data-related assets that are entered into the data inventory or catalog. They should also manage automated data classification, tagging, documentation, and lineage mapping.

The automated data inventory or catalog platform should align with the agency's overall data governance and compliance needs and its data management goals and objectives. It is essential to have a well-established enterprise data governance framework with policies, processes, data stewardship roles, and responsibilities in place to govern the usage of the data inventory or catalog platform and ensure data security, privacy, integrity, compliance, and accountability.

APPENDIX B: DATA STEWARDSHIP

B.1: ESTABLISHING AND MATURING DATA STEWARDSHIP

Data stewardship is the role of individuals or teams across an agency to oversee the management, quality, and security of their agency's data assets. Developing and maturing an enterprise-wide data stewardship program is critical to the success of data categorization and ZT data security. Data stewardship involves the careful management and governance of data throughout its lifecycle, ensuring that data is accurate, accessible, and secure. By establishing clear roles and responsibilities for data stewards, agencies can ensure that data is consistently categorized, maintained, and protected according to defined standards.

Data governance bodies (DGBs) are essential to establishing and overseeing data stewardship programs within their respective agencies as they are required by OMB M-19-23 to set agency data policy and coordinate and support implementation of data management responsibilities with data-management actors within their respective agencies. DGBs guide the establishment of the agency's overarching strategy and framework for data governance, which includes the role of stewardship. By setting clear policies and standards, DGBs facilitate a structured approach to managing data, which is crucial for maintaining data quality, enhancing transparency, and ensuring compliance with regulatory requirements. This strategic oversight ensures that data is leveraged effectively to support decision-making and operational needs while safeguarding privacy and security.

In addition to strategy development, DGBs should provide critical oversight to ensure the data stewardship program is effectively implemented across the organization. They establish accountability by defining roles and responsibilities for data stewards who manage and categorize data assets. DGBs should also monitor the progress and effectiveness of data stewardship activities, adjusting as necessary to address emerging challenges or opportunities. By fostering a culture of data literacy and collaboration, DGBs empower data stewards to uphold best practices in data management and drive continuous improvement in how data is utilized across the agency. This comprehensive approach helps agencies maximize the value of their data assets while minimizing risks associated with data mismanagement.

Data stewards are typically selected from program offices who possess critical domain-specific or business knowledge about their data, making them uniquely qualified to oversee its management. Their deep understanding of the data's context, usage, and relevance allows them to catalog and classify it within data governance tools accurately. By leveraging their expertise, these stewards can ensure that data is properly documented, enhancing its discoverability and usability across the organization. This localized management, supported by an overarching enterprise-wide data governance policy, not only improves data quality and integrity but also aligns data governance practices with the specific operational needs of each program office, fostering a more cohesive and effective data governance framework.

B.2: TYPES OF DATA STEWARDS

When establishing a data steward program, it is beneficial to consider the creation of three distinct roles: technical, business, and information data stewards.

Technical data stewards are responsible for the more technical aspects of data management. Their primary focus is the data infrastructure, including data integration, quality, security, and governance. They ensure data is accurately and efficiently stored, processed, and maintained within the IT systems. Technical data stewards are typically well-versed in database management, data architecture, and IT policies, including a deep understanding of the FISMA risk levels associated with their systems. They work closely with IT departments to implement and uphold data standards and best practices.

On the other hand, **business data stewards** are responsible for the practical application and interpretation of data within the context of business operations. They ensure that the data meets the business needs, is accessible and usable for decision-making, and aligns with business goals. Business data stewards work closely with agency program offices and are typically experts in the specific business domain they support, understanding the data requirements, definitions, and usage within that area.

Meanwhile, **information data stewards** oversee unstructured data — data without a predefined model or organization, such as emails, documents, and multimedia files — ensuring they are categorized by data sensitivity and efficiently managed and disposed of according to established records management policies. By collaborating closely with the Senior Agency Official for Records Management, SAOP, and the Agency Records Officer, information data stewards ensure their program offices adhere to relevant records management statutes, regulations, and policies issued by NARA and OMB.

Identifying these three types of data stewards for their respective data assets is critical for advancing ZT. This data stewardship model ensures comprehensive data management, where both security and business requirements are met, enhancing the overall integrity and trustworthiness of the organization's data environment.

B.3: BUSINESS AND INFORMATION DATA STEWARDS' ROLES IN DATA CATEGORIZATION AND METADATA COMPLETION

Business and information data stewards are essential for effectively categorizing structured and unstructured data, a critical component in maturing ZT. Their expertise and in-depth knowledge of the data they manage enables precise categorization based on sensitivity, criticality, and usage. This precise categorization is vital for ZT environments, where stringent access controls are enforced, and trust is continuously verified. Accurate data categorization by stewards ensures that appropriate security measures are implemented, allowing only authorized individuals to access specific data, enhancing overall data security, and reducing risk.

It is critical that business data stewards categorize their data within data governance tools. Their intimate understanding of the data's context and relevance within the organization ensures that the categorization is both accurate and meaningful. By directly engaging

in the categorization process, data stewards help maintain strict oversight and control over data access and usage, fostering trust and accountability within the organization. Equally important, information data stewards collaborate closely with program office personnel to ensure that all sensitive unstructured data generated by their program office is appropriately labeled, categorized, and protected under their agency-wide CUI and records management policies.

These data stewards play a pivotal role in completing and maintaining metadata, which is essential for effective data governance and implementing ZT. They ensure that metadata accurately describes data assets, such as defining data types, formats, sources, usage contexts, and access permissions. By meticulously cataloging this information using data governance tools, data stewards enhance data's discoverability, usability, and management across the organization. Their domain-specific knowledge allows them to provide precise and relevant metadata that aligns with their respective departments' operational needs and compliance requirements, thereby fostering a comprehensive and reliable data ecosystem.

When considering metadata completion in the context of ZT for a data steward program, several key considerations come into play:

1. Data stewards must ensure that metadata includes detailed access control information, which may include information on who can access the data, under what conditions, and for what purposes.
2. Metadata should capture the data's sensitivity and categorization levels to inform encryption and other security measures.
3. Data stewards must regularly update and audit metadata to reflect changes in data usage or regulatory requirements, ensuring that security policies remain robust and adaptive to emerging threats.
4. The NARA baseline metadata requirements, found in CFR 1236.54, should be considered a baseline for metadata. Other metadata may be captured as needed.

By addressing these considerations, data stewards help create a secure, transparent, and resilient data environment that aligns with ZT principles. Furthermore, as data stewardship matures, it facilitates the integration of advanced technologies and methodologies that bolster ZT security measures. For instance, through data stewardship, organizations can leverage automated tools for data categorization and encryption, ensuring that sensitive data is always protected. By continuously refining data management practices and incorporating feedback from data stewards, organizations can adapt to emerging threats and evolving regulatory requirements. This iterative process strengthens the security posture and ensures that data assets remain resilient against potential threats.

APPENDIX C: SECURITY MONITORING AND CONTROLS

C.1: EXAMPLE IMPLEMENTATION PLAN

Step 1: Initial Setup

- Working alongside the agency data and privacy practitioners, start by choosing a handful of sensitive data asset categories for monitoring.
- Configure system logs to record external data movement to identify if critical data is being used in a risky or unapproved manner, such as being exfiltrated.
- Review logs for accuracy and necessary follow-up actions.

Step 2: Enhanced Actions

- Set up notifications, alerts, quarantines, and blocking for the chosen data asset categories.

Step 3: Expand Coverage

- Include additional data asset categories in the monitoring system.
- Begin with logging and review before implementing further actions.

Step 4: Comprehensive Controls

- Apply the full range of security actions to the new data asset categories.

Step 5: Internal Data Flow Management

- Identify and control internal data transfers within the agency.
- Start with logging and analysis then proceed to more proactive measures.

Step 6: Continuous Improvement

- Implement controls at various points within the network and endpoints.
- Adjust actions as needed to align with mission objectives.
- Assess effectiveness of controls.

Step 7: Iterate and Expand

- Iterate with new criteria and scope.
- Expand to add new sensitive data asset categories for monitoring.

APPENDIX D: ROLES IN DATA SECURITY RISK MANAGEMENT

Table 3 provides examples of roles, responsibilities, and guidance for data security risk management.

As noted in NIST SP-800-37 Appendix D (Roles and Responsibilities), organizations have varying missions, business functions, and organizational structures. Therefore, there may be subtle differences in naming conventions for risk management roles and how risk management responsibilities are allocated among organizational personnel (e.g., multiple individuals filling a single role or one individual filling multiple roles). However, the basic functions should remain the same.

TABLE 3: Roles in Data Security Risk Management

	Who		Does What
	Role	**Data Role**	**Representative Responsibilities and Guidance**
Level 1: Organization	Chief Information Officer (CIO)	Data Owner	The CIO is the organizational officer responsible for establishing information security programs, allocating resources, and ensuring that risk management policies are integrated with organizational goals. The CIO is responsible for the overarching data governance strategy, ensuring that data security and privacy policies are enforced across all levels and throughout the organization.
	Chief Data Officer (CDO)	Data Owner	The CDO enables data search and discovery of data assets with the right context. The CDO also ensures data quality, integrity, and reliability.
	Agency Records Officer (ARO)	Data Custodian	AROs are operationally responsible for their agency's records management program. Federal Records Officers manage and direct records management activities for their agencies to ensure compliance with the Federal Records Act.
	Senior Agency Information Security Officers (SAISO)	Data Custodian	The SAISO develops and maintains the organization's risk management strategy while ensuring alignment with Federal and organizational policies and standards. They guide security control implementation and report risk levels to the executive leadership teams. The SAISOs ensure that data-centric risks are addressed through comprehensive security controls and that all data handling processes comply with relevant laws and policies.

Level 1: Organization	Chief Risk Officer (CRO)	Data Owner	The CRO ensures that risks are consistently identified, assessed, and mitigated across the organization and that there is clear communication between the various levels of the organization regarding risk management efforts. The CRO oversees the identification and management of data-related risks, ensuring that any potential threats to data's confidentiality, integrity, and availability are continuously monitored and mitigated across the organization.
	Executive Leadership (e.g., CEO, Agency Head, other C-suite, etc.)	Data Steward	Executive leadership ensures that the overall organizational risks are managed effectively and efficiently. This is achieved by fostering a culture of security awareness and training — a crucial step in the current data security landscape — to ensure risk and data resources are properly allocated and making informed decisions about risk tolerance and acceptance. Executive leadership sets the tone for data security and privacy within the organization, ensuring that data protection is a priority.
Level 2: Mission/Business Process	Mission or Business Owners	Data Steward	Mission and business owners are responsible for ensuring that secure systems support their operations and data that is critical to their mission is protected. They work closely across various organizational levels to ensure that risk and data management practices are in place and mission-critical functions are protected from threats. They are responsible for overseeing the enforcement of data security measures while maintaining data privacy within their operational areas. For example, this may be represented by a vertical mission capability such as a **Mission Essential Function (MEF)** or crosscutting through various operations areas, such as financial and cyber.
	Bureau/Division Leaders	Data Steward	Bureau and division leaders ensure that security and privacy measures align with mission goals. They are responsible for ensuring that risks within their domain are appropriately identified, mitigated, and communicated to higher organizational levels.
	Security/Privacy Managers	Data Processor, Data Custodian	ZT requires a consolidated view of mission and business data and metadata about the infrastructure (e.g., log data, endpoint data, and privacy). Security and privacy managers are the backbone of an organization's security and privacy. They ensure security and privacy policies and practices are implemented within the bureau's mission or business area. They are responsible for translating organization-level policies into controls.

Level 3: Information System	System Administrators	Data Processor, Data Custodian	This role involves implementing and maintaining controls on systems. System administrators ensure systems are protected with appropriate controls and identified vulnerabilities are addressed.
	Control Assessors	Data Processor, Data Custodian (Support)	Control assessors conduct assessments to verify that controls are correctly implemented and function as intended. Assessment findings are reported to system owners and authorizing officials for risk review, risk response, and system authorization to operate.
	System Owners	Data Controller, Data Owner	System owners are responsible for the overall security and risk management of systems throughout the system development lifecycle, supporting the organization's mission. The system owners ensure that controls are implemented and remain effective for systems under their control. They manage risks, monitor system performance, and report risk information to the organization. System owners play a crucial role in the creation and maintenance of system plans following organizational policies. They are also responsible for ensuring risk assessments are conducted and for responding to risk findings.
	Security and Privacy Officials	Data Processor, Data Custodian	Security and privacy officials are responsible for ensuring that daily operations follow security and privacy policies and procedures while maintaining the security and privacy posture of the system. They conduct regular system monitoring, support incident response, and report risks.

D.1: DATA ROLES AND RISK MANAGEMENT FRAMEWORK ROLES

Table 4 provides a mapping between risk management roles as defined in the NIST RMF and roles required to manage data security risks. Understanding data roles is important for data security, privacy, and risk management practitioners because it ensures that they can collectively and correctly identify responsibilities and accountability for data protection within their organization.

TABLE 4: Mapping Between Data Roles and Risk Management Roles

Data Role	Definition	NIST RMF Role(s)	NIST Reference
Data Subject	The person or entity that information is about.	While not explicitly defined in the NIST RMF, this may align with the conception references to "individuals" or "persons" whose data is collected.	N/A
Data Owner	The entity that collects or creates the data and is responsible and accountable for protecting it.	Chief Information Officer (CIO), Chief Data Officer (CDO), Chief Risk Officer (CRO), Information Owner/Steward, System Owner	NIST RMF, Appendix D
Data Controller	The entity that determines the purpose and meaning of the processing data, as well as ensuring its protection and privacy.	Information Owner/Steward, System Owner	NIST RMF, Appendix D
Data Processor	An entity that works under the direction of the owner/controller, such as an IT department, and processes data in accordance with the instructions of the data controllers and security standards.	Security/Privacy Managers, System Administrator, Security and Privacy Officers, Control Assessors	NIST RMF, Appendix D
Data Custodian	The person or entity responsible for the maintenance and care of data or data sources. For example, implementing technical controls, procedures, and systems.	Senior Agency Information Security Officers (SAISO), Security/Privacy Managers, System Administrator, Security and Privacy Officers, Control Assessors	NIST RMF, Appendix D
Data Steward	Users of the data for mission or business purposes, ensuring data quality and policy adherence.	Executive Leadership (e.g., CEO, Agency Head, other C-suite, etc.), Mission/ Business Owner, Bureau/Division Leaders	NIST RMF, Appendix D

APPENDIX E: IDENTITY, CREDENTIAL, AND ACCESS MANAGEMENT PRINCIPLES

For each of the core elements of **Identity, Credential, and Access Management (ICAM)**, data owners and operators should be aware of the essential practices and understand the underlying reason for those practices. Where applicable, each practice references important external sources of information to which the data owner or operator can consider for further information.

Identity Validation and Verification

Protection of data requires that all entities that are authorized to access data are properly identified and verified. Data owners and operators must ensure that the process of identifying users is performed in such a manner that the level of identity assurance and the level of authentication assurance are reliable and that the access granted to the user is appropriated to risk associated with the strength of the **identity assurance level (IAL)** and **authenticator assurance level (AAL)**. For public-facing users, agencies should use Digital Identity Risk Assessments, as described in NIST Special Publication 800-63,[5] to assess risk and select controls that balance security with access and usability. To understand the level of identity proofing for a given application, refer to the FICAM Digital Identity Risk Assessment Playbook.[6]

Principle of Least Privilege

Data owners and operators should conduct regular access reviews to ensure that users have only the necessary access privileges for their roles. It is the responsibility of the data owner to ensure that there is separation of duties between those that operate systems and those that back them up. This ensures that a compromise of operational accounts does not also compromise backup accounts and that they are available for recovery efforts.

For critical systems deployed in cloud services, data owners should utilize entitlement analytics, such as **cloud infrastructure entitlement management (CIEM)** tools, to provide a holistic view of permissions (entitlements) in cloud environments. Data owners have a responsibility to understand who has entitlement to systems under their control. Cloud services that operate through **application programming interfaces (APIs)** make it difficult to see the totality of permissions and it is an essential practice for critical data systems to be able to view entitlements across the cloud infrastructure to prevent unexpected lateral movement of adversaries who compromise an account.

[5] NIST SP 800-63, *Digital Identity Guidelines*, https://pages.nist.gov/800-63-3/sp800-63b.html

[6] Digital Identity Risk Assessment (DIRA), U.S. Agency for International Development, https://www.usaid.gov/digitalstrategy/dira

Data owners should assess job roles and responsibilities to determine the minimum access privileges required for each role and implement RBAC to assign access privileges to those predefined roles with appropriate responsibilities. This can be a complex and resource-intensive process and should utilize modern machine learning-assisted role analysis tools where available.

User Provisioning and Lifecycle Management

Data owners and operators should separate administrator accounts from user accounts to ensure only designated admin accounts are used for admin purposes.[7] If an individual user needs administrative rights over their workstation, use a separate account that does not have administrative access to other hosts, such as servers. Data owners and operators should use ICAM provisioning services with provisioning accounts or keys to enforce separation of duties based upon criticality of the control. Care should be given as this strategy introduces additional management overhead and is not appropriate in all environments.

Continuous Monitoring

Data owners and operators should utilize **user behavior analytics (UBA)** using either SIEM-integrated modules or through standalone capabilities when these are implemented as part of the Agency ICAM service. UBA can identify patterns and anomalies that may indicate unauthorized access or suspicious activity. Conditional access policies require users who want to access a resource to complete an action. Conditional access policies also account for common signals, such as user or group memberships, IP location information, device, application, and risky sign-in behavior identified through integration using UBA. Refer to the CISA Joint Cybersecurity Advisory *Threat Actor Leverages Compromised Account of Former Employee to Access State Government Organization*, which describes a common attack utilizing a compromised account.[8]

When implementing logging and monitoring mechanisms, data owners and operators should consider how to monitor access in such a manner as to detect potential security events. Refer to the joint CISA and NSA publication *Recommended Best Practices for Administrators*[9] section on Preparation for Implementing Best Practice, which describes key considerations for assessing an organization's logging and monitoring capability to determine which improvements are necessary to counter top threats.

[7] CISA CPG Cross-Sector Cybersecurity Performance Goals, https://www.cisa.gov/sites/default/files/2023-03/CISA_CPG_REPORT_v1.0.1_FINAL.pdf

[8] CISA Threat Actor Leverages Compromised Account of Former Employee to Access State Government Organization, https://www.cisa.gov/sites/default/files/2024-02/aa24-046a-threat-actor-leverages-compromised%20account-of%20former-employee.pdf

[9] CISA Recommended Best Practices for Administrators: Identity and Access Management, https://www.cisa.gov/sites/default/files/2023-12/ESF%20IDENTITY%20AND%20ACCESS%20MANAGEMENT%20RECOMMENDED%20BEST%20PRACTICES%20FOR%20ADMINISTRATORS%20PP-23-0248_508C.pdf

Authentication

Implement continuous authentication mechanisms, such as adaptive authentication or risk-based authentication, to verify user identities throughout their sessions. Do layered authentication policies based on what your access management tool supports or the features of your identity architecture. Integrate MFA mechanisms to match the risk level of access. Utilize identity management systems that automate the identity verification process and provide real-time validation of user identities.

Access Control Mechanisms

Access control mechanisms are essential to the ICAM capabilities in ZT, and they traditionally have various means of implementation. RBAC defines access privileges based on job roles and responsibilities, while ABAC is used for more granular and dynamic access control based on user attributes and contextual information.

Data owners and operators should understand who in their charge should has have elevated privileges, ensuring that they use the principle of least privilege to limit user account permissions to those that are necessary to perform their job. ICAM services often offer **privileged access management (PAM)** tools to control and monitor access to privileged accounts and ensure accountability. The Enduring Security Framework has provided extensive guidance regarding best practices for administrators who are the most common users to utilize privileged access management tools.[10]

For those data owners and operators who utilize RBAC systems that map roles to specific access privileges and dynamically adjust them based on changes in job roles or responsibilities, they can refer to an ISACA resource[11] that helps them understand how to establish appropriate **segregation of duties (SoD)**. Some products can help identify where SoD should be applied but is not applied by using "role mining" techniques. Data owners and operators should ask if their organization's ICAM capabilities offer support for role mining if they are having difficulty understanding how to create appropriate roles.

For data owners and operators that utilize ABAC mechanisms, they can consider attributes like user location, time of access, and other factors to determine access privileges so long as the attributes are available at the time the decision is made. NSA and CISA together have published a guide for securing data in the cloud[12] in which they state, "Consider utilizing a data tagging system or solution, where data is conditionally accessed via granular ABAC policies to protect data. It is also important to separate accounts that grant access to resources from those that manage them daily."

Data owners and operators should use the ICAM capabilities that best match the data that they are trying to protect. CISA has observed that access control decisions are best made using the context in which the access request is being made and refers to CBAC which

[10] CISA Recommended Best Practices for Administrators: Identity and Access Management, https://www.cisa.gov/sites/default/files/2023-12/ESF%20IDENTITY%20AND%20ACCESS%20MANAGEMENT%20RECOMMENDED%20BEST%20PRACTICES%20FOR%20ADMINISTRATORS%20PP-23-0248_508C.pdf

[11] A Step-by-Step SoD Implementation Guide, ISACA, https://www.isaca.org/resources/isaca-journal/issues/2022/volume-5/a-step-by-step-sod-implementation-guide

[12] CISA Secure Data in the Cloud, https://media.defense.gov/2024/Mar/07/2003407862/-1/-1/0/CSI-CloudTop10-Secure-Data.PDF

combines features of RBAC and ABAC to apply dynamic access policies using device-level signals as cues. For more information about CBAC, refer to *Secure Cloud Business Applications: Hybrid Identity Solutions Guidance*.[13]

PAM

Data owners and operators must recognize that privileged accounts require additional monitoring and control over normal users. They should identify which administrators are granted elevated privileges and should be separately managed using a PAM solution with strong identity governance in order to properly manage those users.

Data owners and operators should continue to take advantage of the tools that have been provided by CISA's CDM program that provide PAM in the manner described in the NIST Cybersecurity White Paper. Refer to the CISA: Tool Secures Privileged Access Management report for more information on how to transition from disparate information systems into a cohesive enterprise-wide approach.[14]

Integration with ZT

Data owners and operators should look for opportunities to integrate ICAM control with other components of ZT, such as network segmentation and micro-segmentation, to enforce data protection at every layer. CISA has published guidance on *Layering Network Security Segmentation*[15] that emphasizes the importance of implementing network segmentation where each subnetwork acts on its own. Data owners and operators who operate within cloud services should pay particular attention to the joint NSA and CISA guidance *Use Secure Cloud Identity and Access Management Practices*,[16] which discusses a specific threat model and associated adversary tactics and techniques. Data owners and operators should regularly assess and update access policies and controls to align with evolving ZT requirements and emerging threats.

Governance

Data owners and operators should establish or work within an already established governance framework that includes regular audits, compliance monitoring, and incident response planning to maintain the effectiveness of ICAM control. Refer to *OMB M-19-17*[17] section on "Shifting the Operating Model beyond the Perimeter" when developing a governance structure within the agency. The structure outlined in this chapter identifies a broad group of executives that operate in support of enterprise risk management to effectively drive ICAM efforts.

---

[13] Secure Cloud Business Applications: Hybrid Identity Solutions Guidance, https://www.cisa.gov/resources-tools/resources/secure-cloud-business-applications-hybrid-identity-solutions-guidance

[14] CISA Tool Secures Privileged Access Management, https://www.cisa.gov/sites/default/files/publications/CDM%2520Success%2520Story-CISA%2520PAM%2520Tool%2520.pdf

[15] CISA Layering Network Security through Segmentation, https://www.cisa.gov/sites/default/files/publications/layering-network-security-segmentation_infographic_508_0.pdf

[16] Use Secure Cloud Identity and Access Management Processes, https://media.defense.gov/2024/Mar/07/2003407866/-1/-1/0/CSI-CloudTop10-Identity-Access-Management.PDF

[17] OMB M-19-17: *Enabling Mission Delivery through Improved Identity, Credential, and Access Management*, https://www.whitehouse.gov/wp-content/uploads/2019/05/M-19-17.pdf


federal
CDO
council
CISO
Council
CIO.GOV