NIST JOINT FRAMEWORKS DATA GOVERNANCE AND MANAGEMENT PROFILE CONCEPT PAPER

June 18, 2024

Note to Reviewers

This Concept Paper supports the development of a joint NIST frameworks Data Governance and Management (DGM) Profile. It introduces the basic approach to the DGM Profile, using three illustrative examples, but it is not intended to be the complete Profile.

NIST welcomes all feedback on the concepts within this paper and is interested in answers to the following questions:

- Is the approach of mapping the NIST frameworks to data governance objectives and activities presented in this concept paper helpful? Does it effectively support use of NIST risk management frameworks together? If not, how can it be improved?
- Should the four data governance objectives highlighted in this paper be included in the DGM Profile? Are there other data governance objectives that should be included?
- Should the three data governance and management activities highlighted in this paper be included in the DGM Profile? Are there other activities that should be included?

1. Introduction

The NIST Privacy Engineering Program has received informal stakeholder feedback indicating a desire for resources to support use of the NIST Privacy Framework (PF), Cybersecurity Framework (CSF) and AI Risk Management Framework (AI RMF) together. NIST recognizes that organizations often view privacy, cybersecurity, and AI risk through the lens of data governance and data management. Data governance provides an organizing logic through which authority and control over data management can be exercised.[1] Effective data governance and management supports organizations seeking to leverage data for the development and deployment of innovative systems, products, and services while managing associated risks to privacy, cybersecurity, and AI. These risks are context-dependent and may implicate overlapping organizational priorities. For these reasons, data governance is a helpful starting point for building a joint NIST frameworks resource.

Stakeholders have also described a lack of uniformity around data governance practices. Some organizations take an ad hoc approach, facing challenges such as inconsistent processes or unclear delegation of responsibilities. Organizations with more well-developed data governance may still face challenges when privacy, cybersecurity, and AI domains are “siloed” or excluded from risk management strategy or practices.

The DGM Profile seeks to address these challenges and offer a means to effectively demonstrate complementary use of NIST frameworks and resources. Standards for data governance exist and are used in practice. ISO/IEC 38505-1, for example, offers helpful principles, definitions, and a

[1] *See, e.g.*, Data Management Association. (2009). *DAMA-DMBOK: Data Management Body of Knowledge* (Technics Publications, LLC., New Jersey) at 19.

model for organizational governing bodies to employ.[2] Since the DGM Profile is intended to support organizations using the PF, CSF, and AI RMF together, it focuses on articulating the relationship and dependencies among data governance objectives; data governance and management activities; and privacy, cybersecurity, and AI risk domains.

This paper will support discussion sessions at the NIST public workshop, Ready, Set, Update! Privacy Framework 1.1 + Data Governance and Management Profile Workshop. If you would like to provide informal feedback on this material in addition to or in lieu of participating in the workshop, please send it to privacyframework@nist.gov by July 31, 2024.

More information on the DGM Profile development process can be found in the New Projects section of the NIST Privacy Framework website.

2. DGM Profile Conceptual Approach

This concept paper proposes to create a matrix of general objectives within the domain of data governance and more specific data governance and management activities against which Categories or Subcategories from the three NIST frameworks can be mapped. Organizations can then prioritize these Categories or Subcategories for their specific operational environment or layer on specific Community of Interest Profiles.[3]

2.1. Data Governance Objectives

This concept paper proposes four initial data governance objectives. NIST is interested in whether these are the correct objectives for data governance, whether they should be changed, or whether additional objectives are needed.

- Data quality – Data quality is critical to ensuring that organizational data assets are fit for purpose. Poor data quality can negatively impact an organization's ability to manage cybersecurity, privacy and AI risks. Organizations face numerous challenges managing data quality, and these challenges can be amplified when data comes from multiple sources, is multi-modal, or has uncertain provenance and lineage. Example data quality factors include accuracy, bias, timeliness, completeness, relevance, and consistency.
- Data ethics – Although there is no single definition of data ethics it encompasses the establishment of behavioral norms and standards that support responsible data governance. Ethical standards or norms promote appropriate judgements and accountability at each stage of data processing, from collection through disposal.[4] As noted

[2] *See*, International Organization for Standardization. (2017). *Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data* (ISO/IEC 38505-1:2017(E)).

[3] Examples of Community of Interest Profiles can be found at the National Cybersecurity Center of Excellence Framework Resource Center, see https://www.nccoe.nist.gov/examples-community-profiles.

[4] The NIST Privacy Framework defines data processing as the collective set of data actions (i.e., the complete data life cycle, including, but not limited to collection, retention, logging, generation, transformation, use, disclosure, sharing, transmission, and disposal.
See National Institute of Standards and Technology. (2020). *NIST Privacy Framework: A Tool for Improving Privacy Through Enterprise Risk Management, Version 1.0* (National Institute of Standards and Technology, Gaithersburg, MD) at 29. https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf

in the NIST Privacy Framework, “Although there is no objective standard for ethical decision-making, it is grounded in the norms, values, and legal expectations in a given society.”[5]

- Accountability – Accountability is demonstrated when decision-makers accept responsibility for current and future decisions. Accountable data governance and management for organizations using AI systems involves core concepts of responsible AI, which emphasize human centricity, social responsibility, and sustainability.[6] Accountability is strengthened when organizational decisions about data are supported by rationale and their impacts are understood and documented.
- Data value – Data value may take on different meanings depending on the perspective of different stakeholders, including organizational business/mission value, individual value, and societal value. Value may be found within raw data or from insights gleaned from data analytics. Value may not be apparent or manifest until the organization processes the data or identifies valuable uses (e.g., training an AI model). Tensions may arise when data use creates value for one set of stakeholders to the detriment of other sets of stakeholders. Ethical norms and standards can help organizations manage such tensions appropriately.

2.2. Data Governance and Management Activities

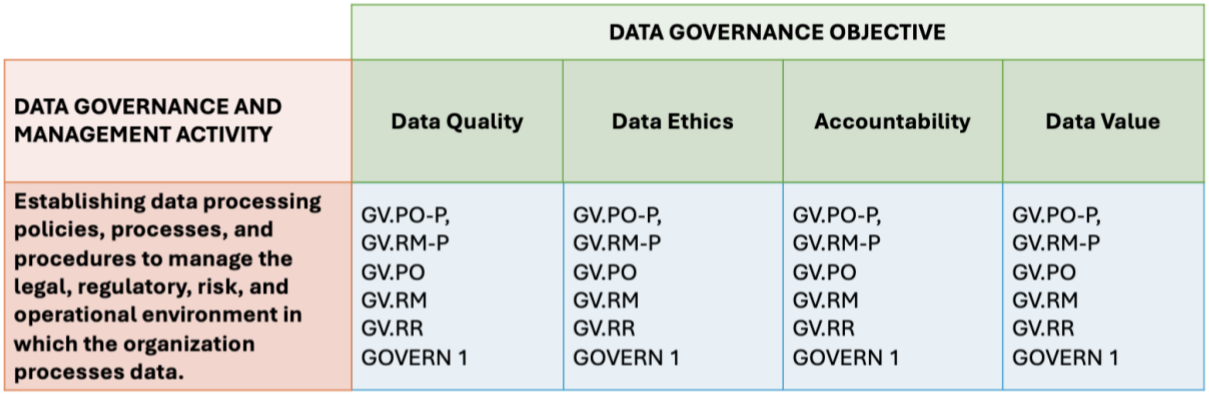
For this concept paper, NIST is proposing the following three data governance and management activities to illustrate how the DGM Profile would work. NIST is interested in whether these are appropriate activities for data governance and management and what other activities should be included in the Profile.

For each activity, NIST has mapped NIST PF, CSF, and AI RMF Categories or Subcategories that can help meet associated data governance objectives. NIST does not include information on Category/Subcategory prioritization in this paper and is interested in whether or how to include prioritization in the final DGM Profile to help organizations tailor their Profile to their unique context, including sector or operating jurisdiction(s). A discussion accompanies each activity to highlight organizational considerations and dependencies among privacy, cybersecurity, and AI risk domains.

2.1.1. Example Data Governance and Management Activity 1: Establishing data processing policies, processes, and procedures to manage the legal, regulatory, risk, and operational environment in which the organization processes data.

[5] *See*, id at 5.

[6] For further discussion on responsible AI, *see* National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (National Institute of Standards and Technology, Gaithersburg, MD) at 1-2. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf



	DATA GOVERNANCE OBJECTIVE			
DATA GOVERNANCE AND MANAGEMENT ACTIVITY	**Data Quality**	**Data Ethics**	**Accountability**	**Data Value**
Establishing data processing policies, processes, and procedures to manage the legal, regulatory, risk, and operational environment in which the organization processes data.	GV.PO-P, GV.RM-P GV.PO GV.RM GV.RR GOVERN 1	GV.PO-P, GV.RM-P GV.PO GV.RM GV.RR GOVERN 1	GV.PO-P, GV.RM-P GV.PO GV.RM GV.RR GOVERN 1	GV.PO-P, GV.RM-P GV.PO GV.RM GV.RR GOVERN 1

PF 1.0 Categories/Subcategories: Governance Policies, Processes, and Procedures (GV.PO-P), Risk Management Strategy (GV.RM-P)
CSF 2.0 Categories/Subcategories: Policy (GV.PO), Risk Management Strategy (GV.RM), Roles, Responsibilities, and Authorities (GV.RR)
AI RMF Categories/Subcategories: GOVERN 1

NIST is interested in whether it would be helpful to provide information about organizational domains that may be implicated by the activity. If so, what organizational domains should be included (e.g., people, process, technology)?

Discussion: Implementing data processing policies, processes, and procedures provides an organized and strategic foundation for data governance and management that addresses privacy, cybersecurity, and AI risks. Organizations will need to tailor their policies, processes, and procedures to their unique context. Factors for consideration include the organization's sector, legal jurisdiction, structure (e.g., functions), and resources.

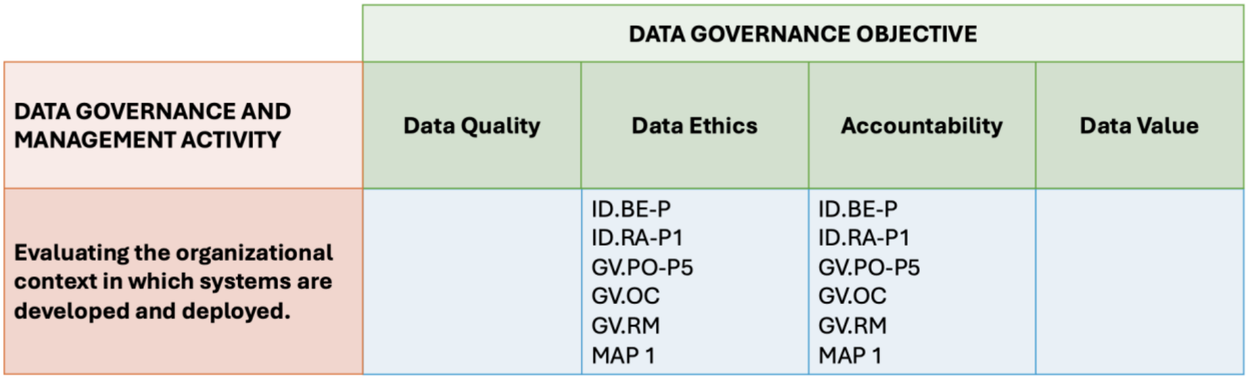
These policies, processes, and procedures support achievement of data governance goals but implementation presents challenges when multiple data processing risks are implicated. For example, establishing the organization's risk tolerance can help strike a balance between maximizing data value while upholding data ethics to guard against negative impacts to individuals, groups, and the organization itself. Risk tolerance is particularly important when prioritizing and responding to privacy, cybersecurity, and AI risks creates tradeoffs (e.g., use of AI systems for real-time threat detection and response may increase efficiency but may introduce problems like incorrect attribution). Establishing and facilitating a cross-organizational understanding of roles and responsibilities for managing privacy, cybersecurity, and AI risks together supports better coordination and accountability for decision-making.

Other policies, processes, and procedures can directly address data quality goals. Effective policies to govern data quality can, for example, address common challenges such as accuracy, completeness, and fairness.[7] The quality of an organization's data assets can affect privacy,

[7] For further discussion, *see, e.g.*, National Institute of Standards and Technology. (2023). *NIST Internal Report 8496: Data Classification Concepts and Considerations for Improving Data Protection, Initial Public Draft* (National Institute of Standards and Technology, Gaithersburg, MD). https://nvlpubs.nist.gov/nistpubs/ir/2023/NIST.IR.8496.ipd.pdf

cybersecurity, and AI risks, particularly when systems require data from multiple sources and across multiple modalities. Such policies can also support data ethics goals by facilitating the application of cybersecurity and privacy controls for organizational data.

2.1.2. Example Data Governance and Management Activity 2: Evaluating the organizational context in which systems are developed and deployed.



	DATA GOVERNANCE OBJECTIVE			
DATA GOVERNANCE AND MANAGEMENT ACTIVITY	**Data Quality**	**Data Ethics**	**Accountability**	**Data Value**
Evaluating the organizational context in which systems are developed and deployed.		ID.BE-P ID.RA-P1 GV.PO-P5 GV.OC GV.RM MAP 1	ID.BE-P ID.RA-P1 GV.PO-P5 GV.OC GV.RM MAP 1	

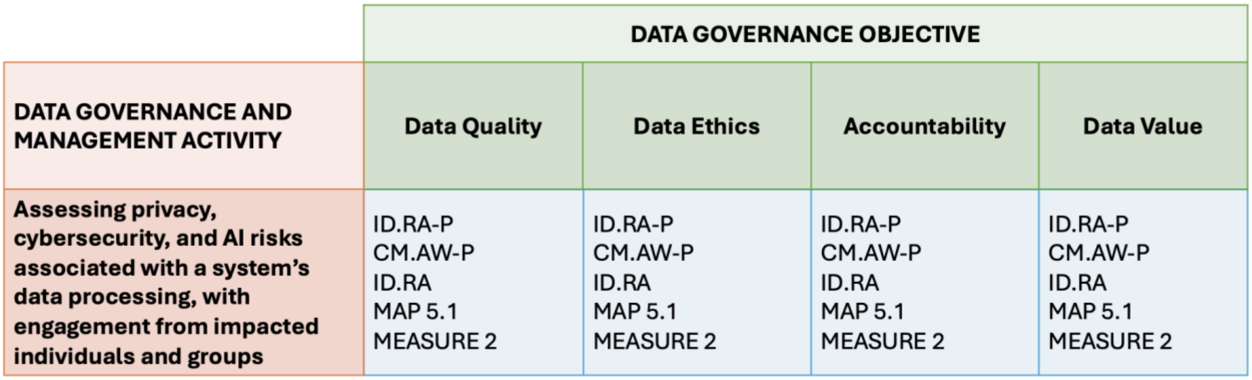
PF 1.0 Categories/Subcategories: Business Environment (ID.BE-P), Risk Assessment Subcategory 1 (ID.RA-P1), GV.PO-P5: Legal, regulatory, and contractual requirements regarding privacy are understood and managed.
CSF 2.0 Categories/Subcategories: Organizational Context (GV.OC), Risk Management Strategy (GV.RM)
AI RMF Categories/Subcategories: MAP 1

Discussion: Systems are developed and deployed within an organization's unique context. These contextual circumstances include the organization's mission, values, risk tolerance, legal and regulatory requirements, role(s) in the data processing ecosystem, and relevant stakeholders. Other contextual considerations such as intended purpose, potentially beneficial uses, and societal norms and expectations gain importance when AI systems are involved.

Organizations can help meet their data ethics goals by identifying and understanding the full context in which a system processes or will process data, including which considerations may need to be weighed heavier than others. For example, privacy, cybersecurity, and AI risks arise in a different context for organizations in highly regulated sectors using AI systems to serve vulnerable communities than for organizations using AI systems with fewer legal requirements and less vulnerable users. Context mapping for AI systems (e.g., defining the boundaries of acceptable deployment) may include contextual factors relevant to privacy and cybersecurity risk assessment such as public perception about the organization's privacy and cybersecurity practices. Thorough evaluation of organizational context in collaboration with broad and diverse interdisciplinary stakeholders can help organizations ensure they are making risk-informed data management decisions consistent with data ethics norms and standards.

The evaluation of organizational context around system design and deployment also supports accountability when system-related concerns are identified and evaluated according to clearly established roles and responsibilities. For example, organizations can document concerns about an AI system's intended purpose within the business context of use and compare it to the organization's stated privacy values, cybersecurity goals, social responsibility commitments, and AI principles. Organizations may need to address identified gaps by reconsidering system design or deployment, including seeking non-AI solutions where costs outweigh benefits.

2.1.3. Example Data Governance and Management Activity 3: Assessing privacy, cybersecurity, and AI risks associated with a system's data processing, with engagement from impacted individuals and groups.



	DATA GOVERNANCE OBJECTIVE			
DATA GOVERNANCE AND MANAGEMENT ACTIVITY	**Data Quality**	**Data Ethics**	**Accountability**	**Data Value**
Assessing privacy, cybersecurity, and AI risks associated with a system's data processing, with engagement from impacted individuals and groups	ID.RA-P CM.AW-P ID.RA MAP 5.1 MEASURE 2	ID.RA-P CM.AW-P ID.RA MAP 5.1 MEASURE 2	ID.RA-P CM.AW-P ID.RA MAP 5.1 MEASURE 2	ID.RA-P CM.AW-P ID.RA MAP 5.1 MEASURE 2

PF 1.0 Categories/Subcategories: Risk Assessment (ID.RA-P), Data Processing Awareness (CM.AW-P)
CSF 2.0 Categories/Subcategories: ID.RA
AI RMF Categories/Subcategories: MAP 5.1, MEASURE 2

Discussion: Regularly evaluating systems for AI, cybersecurity, and privacy risks can help organizations achieve their data governance objectives or identify data processing activities that fall short of data governance goals. For example, an AI system may generate data that is inaccurate or incomplete, creating privacy or safety risks to individuals or groups that are linked to the data, and providing less value than higher-quality data outputs. By assessing this system's risks, the organization can identify data processing activities that fail to meet, or could fail to meet, benchmarks for data quality and data value.

In other cases, AI systems may meet data quality objectives (e.g., by generating accurate data), but such data could be used in unfair or unethical ways. Systems may also generate more data than is necessary to achieve a stated purpose, creating cybersecurity and privacy risks that must be weighed against the organization's goals for data value.

AI system evaluation itself can require data captured from human subjects, introducing risks to the subjects as well as the organization. Regular and sustained engagement with communities impacted by, or potentially impacted by, data processing can help organizations meet accountability objectives and align data processing activities with data ethics goals.