**Department of Homeland Security**

*Response to Federal Evaluation Activities and Data Management (PL 115-435)*

Department of Homeland Security

# Evidence-Based Data Strategy (DHS EDS)

Office of the Chief Information Officer (CIO)
Department of Homeland Security (DHS)

July 16, 2021

# Contents

# Figures

# 1. Executive Summary

Enterprise data strategy and governance are foundational activities for DHS to ensure that the many data management stakeholders across DHS participate in “an integrated and direct connection to evidence needs.”[1] To use the terminology of the Federal Data Strategy 2020 Action Plan,[2] both an “agency data strategy or road map” and “an inclusive and empowered Data Governance Body” are required for DHS compliance with the Evidence Act of 2018.[3] These requirements were reconfirmed as a priority via President Biden’s January 27, 2021 memorandum, “Restoring Trust in Government through Scientific Integrity and Evidence-Based Policymaking.”[4]

This DHS response requires engagement with data stakeholders for their input on data policies, programs, and operations. It emphasizes how data enables progress in meeting DHS strategic objectives and outcomes. DHS will leverage the power of data[5] and enable an evidence-based, data-driven, decision-making environment across the Department.

# 2. Overview

DHS manages significant data holdings across a broad set of missions and activities—many of them public facing and operating in a rapidly evolving threat environment. DHS data are complex, requiring appropriate management, with the assurance of protecting the privacy, civil rights, and civil liberties of the individuals whose information we maintain.

Each of the twenty-two legacy Components of DHS retains ownership and management of its own data, often with decentralized data systems that support the Department’s various operational missions. DHS lacks enterprise-wide data governance to ensure that trusted, critical data are widely available and accessible in a real-time, useable, secure, and linked manner. This limits data-driven decisions and insights in the execution of swift and appropriate action.

The Evidence Act of 2018 provides the opportunity for DHS to establish an enterprise data strategy; moreover, it assigns to DHS the responsibility to promote better use and management of data and evidence consistent with the GPRA Modernization Act and OMB Circular A-11, Part 6.

To meet these requirements, DHS is implementing enterprise-wide data governance to fully address issues that are foundational for strengthening DHS’ ability to support evidence building, leverage information sharing, and promote standards that coincide with DHS-level

[1] Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, OMB Memorandum M-19-23 of July 10, 2019, p. 2

[2] Federal Data Strategy 2020 Action Plan, https://strategy.data.gov/action-plan/

[3] Ibid footnote 1, p.20

[4] “Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking,” Presidential Memorandum, January 27, 2021

[5] Ibid footnote 1, p. 21

strategic planning. DHS data governance will encompass mechanisms and critical data-related criteria, templates, and checklists for setting data standards; validation of data sources, integrity, and best practices; and the completion of agreements across DHS mission areas, management domains, and operational lines of business on information sharing, aggregation of data, and procurement.

## 2.1 Background

The Foundations for Evidence-Based Policymaking Act of 2018, also referred to as the Evidence Act of 2018 (PL 115-435), requires federal agencies to establish processes that modernize data management practices, and thereby better inform policy decisions. This Act builds on longstanding principles from the Government Performance and Results Act (GPRA) Modernization Act that support information quality, access, protection, and use. These principles were updated and further defined in the 2019 OMB update to Circular A-11.[6] Finally, OMB implementation guidance for the Evidence Act establishes a new paradigm that "emphasizes collaboration and coordination to advance data and evidence-building functions….by statutorily mandating Federal evidence-building, open government data, confidential information protection, and statistical efficiency."[7]

The Evidence of Act of 2018 demonstrates that data management stakeholders are embedded in many business disciplines such as policy development, strategic planning, budget development and execution, program administration, performance management, and performance and cost analysis. The Evidence Act of 2018 therefore mandates that agency CDOs implement the needed policies and governance to ensure that their multiple stakeholders participate in "an integrated and direct connection to data and evidence needs."[8]

## 2.2 Mission

Drawing from key themes of the Evidence-Based Policymaking Act of 2018, the new data mission statement of DHS is to:

> *Provide transparent access to valid, reliable, and interoperable data that supports the Department's mission and promotes the public good.* [9]

[6] Integrating Components of the GPRA Modernization Act and Evidence Act to Improve Organizational Performance, https://www.performance.gov/a-11-update/, Performance.gov Team, Aug 3, 2020

[7] Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, OMB Memorandum M-19-23 of July 10, 2019, p.1

[8] Ibid footnote 7, p. 2

[9] DHS data is *valid* when it measures or describes what it claims, and valid data is *reliable* when it measures these claims consistently.

# 3. Guiding Principles

### 3.1 Data-Driven Culture

DHS promotes a data-driven culture that will help employees and contractors value data as a strategic asset. Developing a level of data maturity and appropriate data skills will help the enterprise become more effective in achieving its mission. Plans and programs based on the Department Learning Agenda will promote a data-driven culture through employee assessment, training, and education.[10] Data literacy will be an essential element of the Learning Agenda and will bolster efforts to promote better ownership of data across the enterprise.

- **Invest in Learning:** Promote a culture of continuous and collaborative learning with and about data through ongoing investment in data infrastructure and human resources.
- **Develop Data Leaders:** Cultivate data leadership at all levels of the federal workforce by investing in training and development about the value of data for mission, service, and the public good.
- **Practice Accountability:** Assign responsibility, audit data practices, document and learn from results, and make needed changes.[11]

### 3.2 Ethics-Driven Data Governance

Governance ethics includes the ethical use of data, which will be at the forefront of all plans and actions for how DHS data are collected, used, and shared. Ethical principles regarding the responsible have been and will continue to be important and will be championed by the DHS CDO and all data and analytics leaders across the Department. DHS will "follow appropriate procedures articulated in law and policy when creating, collecting, using, processing, storing, maintaining, disseminating, disclosing, and disposing of data about individuals and entities, whether the data are provided by participants and evaluated entities, or consists of administrative or other data created or obtained from other sources."[12] Component CDOs will be responsible for promoting a culture of ethical data use that is supported by oversight mechanisms which identify and promote best practices among the United States and our partners.[13]

- **Ethics at the Forefront:** DHS will monitor and assess the implications of federal data practices for the public and will design checks and balances to protect and serve the public good.

[10] DMM v.1

[11] Federal Data Strategy 2020 Action Plan, p. 7

[12] Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices, OMB Memorandum M-20-12, March 10, 2020, p. 14-15

[13] Department of Defense Data Strategy, Sep 30, 2020, p. 3. The DHS Chief Data Officer Directorate gratefully acknowledges the Department of Defense (DoD) and its CDO for making available their data strategy document and for encouraging DHS to leverage that work where applicable.

- **Exercise Responsibility:** DHS will practice effective data stewardship and governance and will: employ sound data security practices, protect individual privacy, maintain promised confidentiality, and ensure appropriate access and use.
- **Promote Transparency:** DHS will articulate the purposes and uses of federal data to engender public trust and will document processes and products to inform data providers and users.[14]

### 3.3 Data-Driven Design

The Evidence Act requires development or reconciliation of IT solutions that provide an opportunity to fully automate the information management life cycle; properly secure data; maintain end-to-end management of records; and achieve compliance with applicable statutes, regulations, and other authoritative guidance. The Department will make data management and compliance with policies a top priority of data and system designs. Compliance with required data policies is a critical success factor for continued funding of future DHS solutions and will help determine whether to grant authorizations to operate.[15]

- **Ensure Relevance:** DHS will protect the quality and integrity of its data and will ensure that data are appropriate, accurate, accessible, objective, useful, understandable, visible and timely.
- **Harness Existing Data:** DHS will identify data needs based on priority research and policy questions and will acquire additional data when needed.
- **Anticipate Future Uses:** DHS will create data thoughtfully and will consider its fitness for use by others; will plan for reuse of data from authoritative sources; and will build interoperability into its data from the start.
- **Demonstrate Responsiveness:** DHS will improve its data collection, analysis, and dissemination based on continuous feedback from users and stakeholders in a process that establishes a baseline, gains support, continuously collaborates, and then refines.[16]

### 3.4 Utility of Data

Regardless of the data domain, community, or use, the fundamental challenge is to discover and collect data that can continuously improve the ability of that data to inform decision makers. This process requires increasingly mature practices of data management and data governance.

To achieve this goal, DHS must always enable electronic collection of data at the point of creation and maintain the pedigree of that data. At the moment when data are created, it should be tagged, stored, and cataloged. When data are combined or integrated, the resulting product must likewise be immediately tagged, cataloged, curated, and appropriately secured. To assure that data are visible, the data catalog will be made broadly available

[14] Federal Data Strategy 2020 Action Plan, p. 7

[15] Department of Defense Data Strategy, Sep 30, 2020, p. 4-5

[16] Federal Data Strategy 2020 Action Plan, p. 7

within DHS. To expedite these processes and to minimize the risk of human error, these steps should be automated to the maximum extent possible.[17]

Additionally, Artificial Intelligence (AI), and specifically Machine Learning (ML) algorithms, opens the possibility of collecting and utilizing enormous quantities of data from many data sources to deliver new, breakthrough capabilities. Such opportunities should be considered when DHS evaluates the potential utility of data.

- **Data Analytics:** The above initiatives will facilitate the discovery of useful information within raw data to find trends, answer questions, and support decision making. DHS will apply descriptive, diagnostic, predictive, and prescriptive data analytics tools to maximize the utility of data.

### 3.5 Data Stewardship

Making data available across mission and business systems is essential to gaining an enterprise-wide view into the daily operations of the Department. Making data available is critical to the success of both the Homeland Security Strategy and the Digital Modernization Strategy.

Furthermore, it is U.S. policy to make government data publicly available as permitted within the constraints imposed by privacy and security concerns.[18]

The default posture of DHS is therefore to share information broadly. Data sharing should only be restricted when required by law, by DHS-wide policy, or where security, privacy, or ethical considerations preclude such sharing.

DHS is also responsible for exercising appropriate stewardship of its data so that the consumers of that data will find it to be well-suited for its intended uses. Achieving a mature level of data stewardship will require the assignment of roles throughout the organization, as well as the training of key personnel in the goals and methodologies of data governance and data management.

- **Collective Data Stewardship**: To exploit data fully for decision making, DHS is defining roles and responsibilities for data stewardship. DHS will assign data stewards, data custodians, and a set of functional data managers to achieve accountability throughout the entire data life cycle. DHS will ensure that data stewards establish policies governing data access, use, protection, quality, and dissemination. DHS will ensure that data custodians are responsible for promoting the value of data and enforcing policies, and that functional data managers implement the policies and manage day-to-day data quality.[19]
- **Data Fit for Purpose**: Data that are "*fit for purpose*" are quality data that are readily discoverable and understood within the context of their intended use. This includes careful consideration of any ethical concerns that might arise from the manner in which the data were collected, shared, or used. Special care must be taken to the way that data

[17] Department of Defense Data Strategy, Sep 30, 2020, p. 3

[18] "Open Data Policy – Managing Information as an Asset," OMB Memorandum M-13-13, May 9, 2013.

[19] Department of Defense Data Strategy, Sep 30, 2020, p. 3

are used in visualizations or representations, and the ways that data are integrated, in order to minimize unintended bias. Customers of DHS data have their own requirements for accessing DHS data, which may or may not align with the purpose or intent of the original data collection. Additionally, in some instances, legislation or a regulation may specify how data are to be used and from which source a particular type of data must be consumed. DHS supports data exploration to enhance analyses that will improve decision making.[20]

# 4. Data Goals and Objectives

The Evidence Act of 2018 creates senior-level positions of Chief Data Officer, Evaluation Officer, and Statistical Official, and calls on these newly designated leaders to work together to support the use of evidence and data within the Department. DHS requires these newly designated leaders to coordinate closely with other DHS leaders—such as the Performance Improvement Officer, Chief Information Officer, and others—to fulfill their duties. In practice, this means that DHS requires management, operations, and line of business officials to look beyond their typical partners and consider how others, such as these newly designated officials, can support and enhance their various functions.[21]

DHS will guide Evidence Act implementation via this strategy and via governance processes which ensure that evidence generation becomes integrated into existing agency processes and operations, including performance management activities. DHS will seek to break down legacy functional silos such as evaluation, performance, and strategic planning and replace them with coordinated collaboration and leadership in service of DHS mission delivery.

### 4.1 Goal: Make Data Visible

The DHS measure for data sharing and data visibility is how easy it is for the data to be discovered and used in creating meaningful data analyses that have depth and breadth. From tracking the right data to integrating it with management measures, clarity of data is critical for informed decision making. “DHS is obligated to make data visible to authorized users by identifying, registering, and exposing data in a way that makes it easily discoverable across the enterprise.”[22]

DHS will know it has made progress toward making data visible when:

**Objective 1:** All DHS data assets have been made visible and available for authorized users when and where needed.

**Objective 2:** Metadata standards have been implemented for DHS.

**Objective 3:** All DHS data assets have been cataloged, including location, access methods, and standards compliance.

[20] Department of Defense Data Strategy, Sep 30, 2020, p. 4

[21] Ibid, footnote 5

[22] Department of Defense Data Strategy, Sep 30, 2020, p. 6

**Objective 4:** Common services have been implemented to publish, search, and discover all DHS data assets.

**Objective 5:** DHS Law enforcement and business governance bodies are making decisions based on live visualizations of near-real-time data.

### 4.2 Goal: Make Data Accessible

The DHS measure for data accessibility is its ease of availability to authorized users in the most relevant and meaningful forms, including high-performance data access for analytics and machine learning, and the display of visualizations via the equipment with which users are provided, such as laptops and smartphones. Accessibility encompasses the security controls and other protective mechanisms that are in place for credentialed users, to ensure that access is permitted in accordance with laws, regulations, and policies.[23]

DHS will know it has made progress toward making data accessible when:

**Objective 1:** Data are inventoried in a comprehensive data catalog with relevant information on purpose, ownership, points of contact, security, standards, interfaces, limitations, and restrictions on use.

**Objective 2:** Data are accessible through documented standard Application Programming Interfaces (APIs).

**Objective 3:** Common platforms and services create, retrieve, share, utilize, and manage data.

**Objective 4:** Data access and sharing is controlled through reusable APIs.

**Objective 5:** Visualizations such as data-driven dashboards can be readily created and displayed on DHS-provided laptops and smartphones.

### 4.3 Goal: Make Data Understandable

The DHS measures for understandable data are the quantity and quality of sharable insights and visualizations for decision makers that can be taken from the aggregation, comparison, analysis, interpretation, and contextualization of complex information, data sets, and ideas.

DHS will know it has made progress toward making data understandable when:

**Objective 1:** Data are presented in a way that preserves its semantic meaning and is expressed in a standardized manner.

**Objective 2:** Common data syntax is used for the same data types and semantic metadata are included with data assets.

**Objective 3:** Data elements are aligned into a comprehensive data dictionary such as the National Information Exchange Model (NIEM) with a controlled, yet flexible, vocabulary and taxonomy.

[23] Ibid, footnote 8, p. 7

**Objective 5:** Processes are implemented to create, align, implement, and manage business vocabularies, including enterprise standards.

**Objective 6:** Adaptive, intelligent systems monitor data streams and identify opportunities to transform, combine, or derive new data, thereby providing increased insights.

### 4.4 Goal: Make Data Linked

The DHS measures for linked data are that the Department adheres to industry best-practices for open data standards, data catalogs, and metadata tagging, and ensures that connections across disparate sources, relationships, and dependencies can be uncovered, maintained, and leveraged for analytics.[24]

DHS will know it has made progress toward making data linked when:

**Objective 1:** Unique identifiers have been implemented in a consistent manner across multiple datasets so that data can be easily discovered, linked, retrieved, and referenced.

**Objective 2:** Common metadata standards have been established that allow data to be joined and integrated.

### 4.5 Goal: Make Data Trustworthy

The DHS measures for data trustworthiness are the documentable quantitative and qualitative credibility, transferability, dependability, and confirmability of DHS information for authorized users and stakeholders.[25]

DHS will know it has made progress toward making data trustworthy when:

**Objective 1:** DHS budget requests and the supporting budget process integrate data-focused evidence and Learning Agendas (see P.L. 115-435).

**Objective 2:** Data have associated metadata about protection, lineage, pedigree and quality bound to them throughout their life cycle.

**Objective 3:** Data quality management techniques are executed to assess and enhance data quality.

**Objective 4:** Master Data Management has been implemented for business, intelligence, and law enforcement data.

**Objective 5:** Data and records have been properly tagged and are being maintained in accordance with established processes and policies.

### 4.6 Goal: Make Data Interoperable

The DHS measures for data interoperability are the quality and quantity of machine-to-machine communications across different technology systems and software applications.

[24] Ibid, footnote 9, p. 8

[25] Ibid, footnote 9, p. 8

DHS seeks to use common semantic and syntactic data formats to exchange data accurately, effectively, and consistently. Machine-to-machine communications advance algorithm development and provide a strategic advantage to the Department.[26]

DHS will know it has made progress toward making data interoperable when:

**Objective 1:** Data exchange specifications have been documented and implemented for all systems, including those of coalition partners.

**Objective 2:** Exchange specifications contain required metadata and convey standardized semantic meaning is conveyed along with the data set.

**Objective 3:** Public data assets are machine-readable and available for consumption.

**Objective 4:** Differing data standards and formats are rapidly mediated to ensure compatibility without mission-critical loss of fidelity, precision, or accuracy.

**Objective 5:** A data tagging strategy and subsequent implementation plan have been developed and promulgated to enable data interoperability.

### 4.7 Goal: Make Data Secure

The DHS measures for data security are the degree to which the guiding principles of risk prioritization, cost effectiveness, innovation, agility, and collaboration are being leveraged in applying the Department's five-pillar cybersecurity strategy[27] for the protection of information through risk identification, vulnerability reduction, threat reduction, consequence mitigation, and cybersecurity outcome enablement, in order to foster resiliency across software, hardware, services, and technologies.

DHS will know it has made progress toward making data secure when:

**Objective 1:** Granular privilege management (based on identity, role, or attribute as required) has been implemented to govern the access to, use of, and disposition of data.

**Objective 2:** Data stewards regularly assess classification criteria and test compliance to prevent security issues from occurring as a result of data aggregation.

**Objective 3:** Approved standards have been implemented for security markings, data handling restrictions, and records management.

**Objective 4:** Classification and control markings have been defined and implemented; content and record retention rules have been developed and implemented.

**Objective 5:** Data loss prevention technology has been implemented to prevent unintended release and disclosure of data.

**Objective 6:** Only authorized users can access and share data.

**Objective 7:** Metadata on access and handling restrictions has been bound to data in an immutable manner.

[26] Ibid, footnote 9, p. 8

[27] Cybersecurity Strategy, U.S. Department of Homeland Security, May 15, 2018

**Objective 8:** The access, use, and disposition of data are being fully audited.

# 5. Data Governance Framework

DHS has been collecting, analyzing, managing, and reporting on multiple categories of data for many years. DHS stakeholders use this data to measure progress of mission programs, usage of mission and business resources, and quality of services. Stakeholders have seen substantial achievements across the mission and business domains that have been documented in DHS strategy and performance progress reports.

However, the Department's efforts to leverage, collect, analyze, manage, and report data could be improved, consistent with the DHS Data Framework Act[28] for more effective and timely enterprise outcomes. Sections of this strategy document address the guiding principles and data objectives by which DHS will seek these improvements. DHS will move beyond transparency to (1) leveraging data *internally* to continuously improve policy, processes, and resource allocation, and (2) making data open *externally* for appropriate public use.

## 5.1 Governance-Centered Data Management

OMB implementation of the Evidence Act in M-19-23 states that CDOs shall be in a central leadership position within agencies, so that they can exercise their authority and responsibilities for data governance and life cycle management to enable data-driven decision making.

Within DHS, the CDO's leadership strength is vested in the governance authority and responsibilities of the DHS Data Governance Council (DGC). Chaired by the DHS CDO, authority is exercised by the decisions of the Council, which are achieved with the advice and support of the Evaluation Officer and Statistical Official co-chairs and of the Domain and Component Membership. Together, these officials and Council members collaboratively and continually work to improve the production of evidence for use in DHS policymaking.

[28] Department of Homeland Security Data Framework Act, Public Law 115-331, December 19, 2018

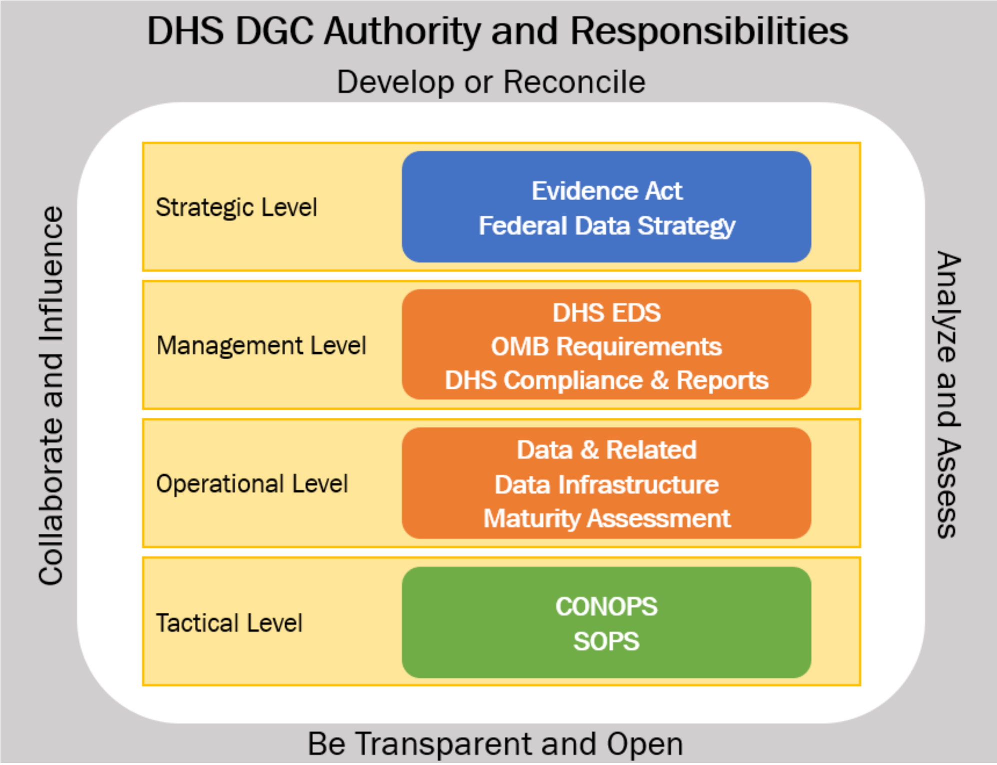

**Figure 1: DHS DGC Authority and its Responsibilities**

Supported by the vested authority of Council decisions, the DHS CDO employs extensive collaboration across related Councils and Boards, and with Mission and Line of Business leadership, as shown in Figure 1, to lead DHS to higher levels of data maturity and improved data management practices.

### 5.2 Governance Across Data Domains

OMB implementation of the Evidence Act in M-19-23 requires a new paradigm by calling on agencies to significantly rethink how they currently plan and organize evidence building, data management, and data access functions, to ensure an integrated and direct connection to needs for data and evidence.[29] The DHS response engages senior leadership to cooperate with the DHS CDO across mission and management data domains to develop evidence building as the basis of the DHS data strategy and to accelerate the governance of DHS data.

The DHS CDO is rolling out a multi-year, collaborative effort with data domain leads, program managers, data managers, and system owners to assess DHS data maturity and data infrastructure maturity across DHS systems. As shown in Figure 2, this assessment will focus on maintaining up-to-date inventories of system data assets and associated data sets and on improving system data cataloging, strategy, processes, protection, efficiencies, and availability. With the insights and findings from these assessments, the DHS Data Governance Council will begin the process of:

[29] Ibid footnote 3, p. 2

- Documenting DHS best practices for the generation, protection, dissemination, sharing, and use of data
- Identifying ways DHS can improve the production of evidence for use in policymaking
- Improving access to DHS data assets
- Coordinating DHS Data Governance Council activities to ensure compliance with OMB guidance

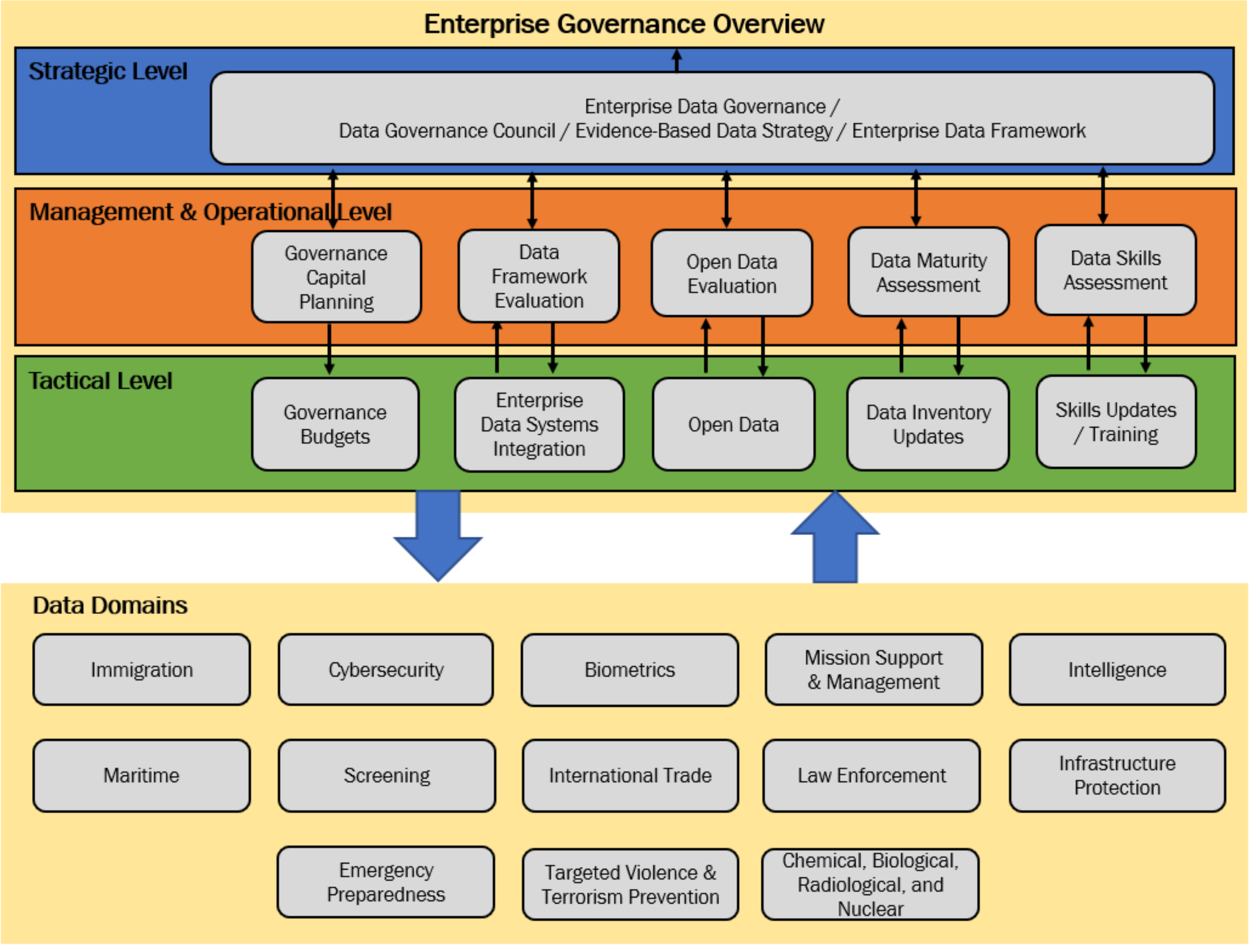
Enterprise Governance Overview
Strategic Level
Enterprise Data Governance / Data Governance Council / Evidence-Based Data Strategy / Enterprise Data Framework
Management & Operational Level
Governance Capital Planning
Data Framework Evaluation
Open Data Evaluation
Data Maturity Assessment
Data Skills Assessment
Tactical Level
Governance Budgets
Enterprise Data Systems Integration
Open Data
Data Inventory Updates
Skills Updates / Training
Data Domains
Immigration
Cybersecurity
Biometrics
Mission Support & Management
Intelligence
Maritime
Screening
International Trade
Law Enforcement
Infrastructure Protection
Emergency Preparedness
Targeted Violence & Terrorism Prevention
Chemical, Biological, Radiological, and Nuclear

**Figure 2: Enterprise Data Governance Overview**[30]

### 5.3 DHS Data Maturity Model

The DHS CDO is responsible to assess the Department's data maturity, data infrastructure, and associated risks, so that improvements can be made in data management processes and priorities for data investment can be recommended.

[30] Figure prepared for the April 2021 DHS Data Governance Council meeting. The number and names of the DHS data domains will change over time.

Figure 3 shows the six categories of data governance and data maturity defined by the DHS Data Maturity Model (DMM). The DMM provides measurable definitions of data maturity, at five different levels of maturity (Level 1 through Level 5), for each of the process areas within the six categories in the model. DHS will use those definitions of maturity as the basis for creating the content of the data maturity assessments required by OMB for implementation of the Evidence Act.

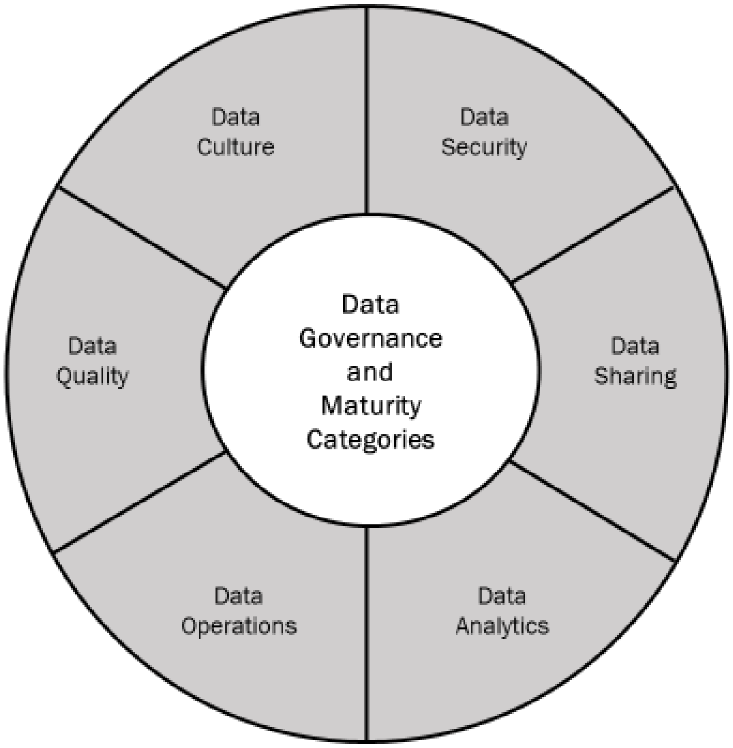

**Figure 3: DHS Data Maturity Model Categories**

**Governance is the most important part of data management. It links together the six categories within this Data Maturity Model and is the glue that ties together the processes within each category, as shown in Figure 3. The DHS CDO will use assessments of data maturity based on the metrics and processes summarized below to help inform and guide the Department's governance and management of its data assets.**

### 5.3.1 **Measure Data Security**

The Data Security category includes processes which address: compliance with existing data security policies, processes, procedures, and training; ensuring that security and integrity are maintained when data are being shared and exchanged; and enabling access controls, business rules, and other security-related metadata that travels along with data as it is being managed, stored, and shared.

The purpose of measuring data security is to quantify and qualify the effectiveness of information security planning and data protection in data system implementation and operations and document the results of data system compliance with their own data security plans and processes.

Data security metrics are established to:

- Confirm appropriate processes for access and change to data assets
- Confirm compliance with regulatory requirements for privacy and confidentiality
- Confirm and maintain data protection for data being shared
- Data security metric procedures are to assess the effectiveness of system data security policy, standards, controls, and procedures:
- Conduct security planning
- Implement security metadata tagging

### 5.3.2 **Measure Data Sharing**

The Data Sharing category includes processes that help build an organizational culture which values data and promotes public use. These processes include making data available to other federal agencies, to state, local, and tribal governments, to non-governmental entities, and to the public. Plans and actions to increase data sharing will also enable richer data analyses and better decision making.

The purpose of measuring data sharing is to document data system sharing plans, events, and outcomes, and identify the resulting benefits and issues.

Data sharing metrics are established to:

- Percent of data assets that are available in JSON and XML formats
- Percent of data assets that have public APIs associated with them
- Percent of data assets that comply with the National Information Exchange Model (NIEM)
- Percent of data assets for which complete metadata are available for download

Data sharing metric procedures are to assess the effectiveness of system data sharing policy, standards, controls, and procedures:

- Make updates to the system and data asset inventory
- Make updates to the Federal Data Catalog
- Implement an Open Data plan
- Implement data sharing initiatives
- Promote data access

### 5.3.3 **Measure Data Analytics**

The Data Analytics category includes processes that support the capacity to meet both current and emerging analytic requirements and challenges. Data Analytics processes include developing necessary skill sets and identifying and preparing required data sets. Processes must ensure that the organization is utilizing both robust business intelligence tools and emerging analytical techniques and technologies. Advanced analytical skills, tools, and techniques also need to be applied to ongoing process improvement efforts throughout the organization

The purpose of conducting data analytics is to increase capabilities for making informed decisions.

Data analytics metrics are established to:

- Understand business intelligence information needs
- Evaluate business intelligence architecture
- Evaluate customer user satisfaction

Data analytics metric procedures are to assess data skills and the effectiveness of an infrastructure that will support all other data management functions and will deliver consistent business intelligence:

- Define and develop data skills
- Conduct analytics and business analytics
- Define and develop infrastructure maturity

### 5.3.4 **Measure Data Operations**

The Data Operations category includes processes which provide critical quality and control functions. These processes ensure that data requirements have been validated and prioritized by stakeholders and are well documented; data are traceable through all of the business and IT processes that produce and consume it; changes to data and its related processes are properly managed; and that data used by business processes satisfies business objectives.

The purpose of measuring data operations is to demonstrate efficiency and value, document customer and stakeholder satisfaction, and measure deviations from specifications.

Data operations metrics are established to:

- Establish measurement and analysis standards
- Define and categorize risks and root causes
- Measure performance and deviation patterns

Data operations metric procedures are to assess the effectiveness of system data operations policy, standards, controls, and procedures:

- Define data management methodology
- Develop requirements and design

- Implement metadata management
- Implement document and content management
- Implement reference data and Master Data Management (MDM)

### 5.3.5 Measure Data Quality

The Data Quality category includes processes that help make data more effective in satisfying the purposes and requirements of its intended use. The measurement of data quality involves assessing such characteristics as its validity, accuracy, completeness, consistency, reliability, availability, accessibility, and timeliness. The achievement of measurable improvements requires the combined efforts of people, processes, and technology throughout the entire data life cycle.

The purpose of measuring data quality is to document the achievement of measurable improvements in the quality of data over time.

Data quality metrics are established to:

- Improve the quality of data in relation to defined business expectations
- Define requirements and specifications for integrating data quality control into the system life cycle
- Define processes for measuring, monitoring, and reporting conformance to acceptable levels of data quality
- Data quality metric procedures are to determine the degree to which data adequately supports the mission and business needs of DHS:
- Develop and manage standards and metrics
- Develop and manage profiling and measurement
- Measure and reduce the amount of duplicative data that contributes to data quality issues
- Measure the timeliness of data
- Measure the amount of data validation being performed
- Develop and manage issue resolution
- Execute data life cycle governance

### 5.3.6 Measure Data Culture

The Data Culture category includes processes that help staff members at all levels of the organization value data as a strategic asset. The organization's workforce requires data awareness and data maturity in order to help the enterprise become more effective in achieving its mission. Plans and programs based on the organization's Learning Agenda promote a data-driven culture through workforce assessment, training, and education.

Business data stewards play critical roles in both the data governance process and in enabling improvements to data quality and other data management process areas.

The purpose of measuring data culture is to identify how data helps the organization enhance performance and measure success:

Data culture metrics are established to:

- Communicate data management and ethics standards, policies, and processes
- Adjust data management standards, policies, and processes based on feedback
- Apply business needs and desired outcomes to data improvements

Data culture metric procedures are to assess the maturity of the organization and users in data awareness, in valuing data as a strategic asset, and in helping the enterprise use data to become more effective in achieving its mission:

- Develop and maintain data literacy
- Foster data ethics and improve the ethical usage of data
- Promote and implement data stewardship

### 5.4 Data Maturity Benchmark

Data maturity is a measurement of DHS' ability to undertake continuous improvement in the six categories of data governance and management defined in the DMM: data security, data sharing, data analytics, data operations, data quality, and data culture. The DHS CDO's next step is to conduct the required data maturity assessments, in order to:

- Assess current maturity levels within the Department
- Identify the desired levels of data maturity
- Evaluate the gaps, and develop plans for addressing them
- Compare the assessment results with peers in other departments and agencies and within industry

The results of the maturity assessments will assist DHS data domain leads, program managers, data managers, and system owners in developing roadmaps, actions, and implementation plans for making measurable progress toward the desired future state of data maturity. Compliance with Evidence Act requirements and OMB guidelines will require continuous improvement efforts over time. A high-level, notional timeline for the maturity assessments and benchmarks is shown in Figure 4.

The FY2021 benchmark for CDOD is expected to focus on the Immigration, Management, and Cybersecurity domains. The FY2022 benchmark is expected to focus on the Emergency Preparedness and Law Enforcement domains. Further implementation plans, training, templates, and processes will be announced from and by the DHS Data Governance Council under the leadership of the DHS CDO.

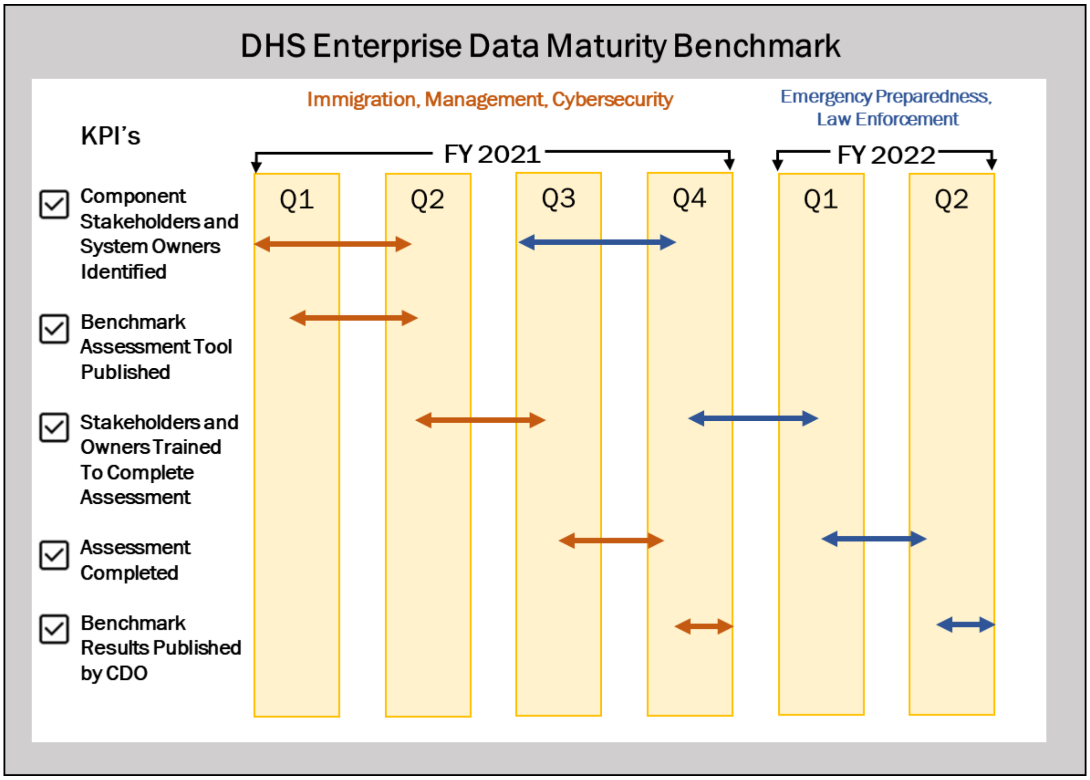

**Figure 4: DHS Data Maturity Benchmark Timeline**

# 6. Conclusions

DHS is facing an urgent need to grow its data-related capabilities and to mature its data management processes, its underlying data infrastructure, and its enterprise data governance processes. With the publication of this Evidence-Based Data Strategy, the Department is at the threshold of conducting a broad assessment of its data maturity and of creating a roadmap that will lead to significant capability improvements in the future. In addition to the data that will be gathered from the maturity assessments themselves, good information gathering will also require frequent engagement with many internal and external data stakeholders. It is essential that DHS gain a god understanding of the challenges faced by those who are already relying on its data, as well as an understanding of the data maturity gaps and specific use cases that currently exist.[31]

As the process of assessing data maturity begins, it is important to realize that the outcome of the assessments will reveal both current issues and current capabilities that may not have been apparent to date. It is critical that the DHS Data Governance Council and the CDO define the desired future state of DHS data maturity by applying the guiding principles, goals, and objectives of this strategy, along with the data maturity measures contained within the DHS Data Maturity Model. A major benefit of considering the desired future state now, at the beginning of the assessment process, is that DHS will be reminded that the

[31] Setting the Foundation for Data Maturity, World Wide Technology, Data Analytics and AI, Kevin Wald and Christopher Graham, https://www.wwt.com/article/setting-the-foundation-for-data-maturity, Mar 30, 2020

purpose of conducting an assessment is to facilitate the Department's growth along the data maturity curve, in order to realize the full benefit of both new and existing DHS data capabilities.[32]

ERIC N
HYSEN

Digitally signed by ERIC
N HYSEN
Date: 2021.07.16
17:37:44 -04'00'

Chief Information Officer

[32] Ibid, footnote 30

# Appendix A. Authorities

*Authorities are presented in alphabetical order based on title.*

"Department of Homeland Security Data Framework Act of 2018," Public Law 115-331, December 19, 2018.

*How applicable*: "This [law] directs the Department of Homeland Security (DHS) to: (1) develop a data framework to integrate existing DHS data sets and systems for access by authorized personnel in a manner consistent with relevant legal authorities and privacy, civil rights, and civil liberties protections; (2) ensure that all information of a DHS Office or Component that falls within the scope of the information sharing environment, and any information or intelligence relevant to priority mission needs and capability requirements of the homeland security enterprise, is included; and (3) ensure that the framework is accessible to DHS employees who have an appropriate security clearance, who are assigned to perform a function that requires access, and who are trained in applicable standards for safeguarding and using such information."

*Why important*: Excluding from this law information that may jeopardize protection of sources, compromise a criminal or national security investigation, duplicate or not serve operational purpose, or be inconsistent with other federal laws or regulations; and including auditing capabilities, mechanisms for identifying insider threats and security risks, and protecting privacy, civil rights, and civil liberties, the DHS Data Framework Act issued "guidance for DHS employees authorized to access and contribute to the framework that enforces a duty to share between DHS Offices and Components for mission needs; to promulgate data standards, and to instruct DHS Components to make available information through the framework in a machine-readable standard format."

Note that the references above are from the summary of H.R. 2454, H.R.2454 - 115th Congress (2017-2018): Department of Homeland Security Data Framework Act of 2018 Congress.gov | Library of Congress

Federal Data Strategy 2020 Action Plan, President's Management Agenda, updated May 14, 2020.

*How applicable*: This action plan provides a common set of data principles, best practices, and practice-related steps for a given year in implementing data innovations that drive more value for the public.

*Why important*: The 2020 action plan establishes a foundation of principles and practices that will support agency implementation of the strategy over the next decade. Specifically, the plan identifies initial actions for agencies that are essential for developing and reconciling processes, building capacity, and aligning existing efforts to better leverage data as a strategic asset.

"Foundations for Evidence-Based Policymaking Act of 2018," Public Law 115-435, 132 Stat. 5529, January 14, 2019.

*How applicable*: This law amends titles 5 and 44, United States Code, to require Federal evaluation activities, to improve Federal data management, and for other purposes.

*Why important*: This law emphasizes collaboration and coordination within and across agencies by statutorily mandating Federal evidence-building activities, open government data, confidential information protection, and statistical efficiency.

Government Performance and Results Act (GPRA) Modernization Act of 2010, Public Law 111-352, January 4, 2011.

*How applicable*: The GPRA Modernization Act improved the Federal Government's performance management framework by increasing the accountability and transparency of program performance and service delivery.

*Why important*: GPRA requires agencies to compare strategy and actual achievements by establishing and measuring progress against performance requirements and goals, operational processes, budgeting strategies, and technology and skill requirements.

Homeland Security Act (HSA), Public Law 107-296, 116 Stat 2135, November 25, 2002.

*How applicable*: Relative to data, the HSA authority can be interpreted to include requirements for a wide variety of protected and personal data that must be protected from unauthorized disclosure. This resulted in the first statutorily required privacy office, responsible for evaluating privacy's effect on the Department's data, which evolved into addressing challenges for the agency in determining how to respond to increasing requirements for data transparency.

*Why important*: The original responsibilities of DHS that include directing and controlling investigations and associated sensitive information have not changed, requiring the newly established position of DHS Chief Data Officer and the CDO organization to balance requirements to increase data transparency for the public with security concerns about exposing data and systems for unlawful use or intrusion. Finding the appropriate balance between the two is the focus that DHS will have in establishing enterprise data governance across departmental mission and operational domains.

Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking, White House Presidential Actions, January 27, 2021.

*How applicable*: This Presidential action establishes the "policy of the Administration is to make evidence-based decisions guided by the best available science and data."

Why important: "Scientific and technological information, data, and evidence are central to the development and iterative improvement of sound policies, and to the delivery of equitable programs, across every area of government."

"Modernizing Government for the 21st Century," President's Management Agenda, March 20, 2018.

> *How applicable*: The 2018 PMA established a cross-agency priority goal focused on leveraging IT modernization and data as a strategic asset, which led to the creation of the Federal Data Strategy.
>
> *Why important*: The 2018 PMA laid out a long-term vision for modernizing the Federal Government in key areas, including modernizing information technology and improving data accountability and transparency to increase agencies' ability to deliver mission outcomes, provide excellent service, and effectively steward taxpayer dollars.

"Managing Information as a Strategic Resource," OMB Circular No. A-130, Appendices I and II, July 28, 2016.

> *How applicable*: "This Circular1 establishes general policy for the planning, budgeting, governance, acquisition, and management of Federal information, personnel, equipment, funds, IT resources, and supporting infrastructure and services.
>
> *Why important:* Appendices I and II to this Circular also include responsibilities for protecting Federal information resources and managing personally identifiable information (PII).

"Making Open and Machine Readable the New Default for Government Information," Executive Order 13642, May 14, 2013.

> *How applicable*: This EO emphasized that openness in government strengthens democracy and promotes the delivery of efficient and effective services to the public.
>
> *Why important*: This EO emphasized the need to make information resources easy to find, accessible, and usable to fuel innovation, discovery, and economic growth.

OMB Memorandum, "Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, M-19-23, July 10, 2019.

> *How applicable*: This memo aligns the requirements of the Foundations for Evidence-Based Policymaking Act of 2018 ("Evidence Act") with and complements components of the Government Performance and Results Act (GPRA) Modernization Act and demonstrates how OMB Circular A-11 provides a framework for thinking through how deliverables relate to one another.
>
> *Why important*: The framework provided by M-19-23 helps agencies strengthen the integration of strategic planning, strategic reviews, and critical management and learning components to advance evidence-building efforts, improve performance, and enable progress in meeting agencies' strategic objectives and outcomes.

"Open Data Policy – Managing Information as an Asset," OMB Memorandum M-13-13, May 9, 2013.

> *How applicable*: Pursuant to EO 13642, this Memorandum required agencies to collect or create information in a way that supports downstream information processing and dissemination activities. This includes using machine readable and open formats, data standards, and common core and extensible metadata for new information creation and collection.
>
> *Why important*: Pursuant to EO 13642, this Memorandum ensured information stewardship using open licenses and review of information for privacy, confidentiality, security, or other restrictions to release. Additionally, it supported IT modernization of information systems to improve interoperability and information accessibility, maintain internal and external data asset inventories, enhance information safeguards, and clarify information management responsibilities.

Part 6: Strengthening the Policy Framework for Improving Program and Service Delivery, OMB Circular A-11, July 10, 2020.

> *How applicable*: OMB's A-11 update calls on agencies to align critical management and learning components of the GPRA Modernization Act and the Foundations for Evidence-Based Policymaking Act of 2018 to support agency leaders in more effective delivery on the missions of their agencies.
>
> *Why important*: "The 2020 A-11 update clarifies how the deliverables required by the Evidence Act of 2018 relate to and mutually reinforce agency strategic planning and other organizational performance and program service planning, delivery, and reporting. This framework will help agencies strengthen the integration of strategic planning, strategic reviews, and Learning Agendas—key organizational learning and planning activities—to advance broad methodological approaches for evidence-based and iterative development and delivery of policies, programs, and operations, which are designed to improve performance and enable progress in meetings agencies' strategic objectives and outcomes."
>
> *Note that a version of the July 10, 2020 update was published in December 2020 without Part 6, although the position of Part 6 in the document was not replaced and the document sections were not renumbered. There is indication that additional information and amplification of this section will be provided by OMB as part of OMB's implementation of the January 27, 2021 White House Memorandum on Restoring Trust in Government through Scientific Integrity and Evidence-Based Policymaking.*

"Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, OMB Memorandum, M-19-23, July 10, 2019.

> *How applicable*: This memo aligns the requirements of the Foundations for Evidence-Based Policymaking Act of 2018 ("Evidence Act") with and complements components of the Government Performance and Results Act (GPRA) Modernization Act and

demonstrates how OMB Circular A-11 provides a framework for thinking through how deliverables relate to one another.

*Why important*: The framework provided by M-19-23 helps agencies strengthen the integration of strategic planning, strategic reviews, and critical management and learning components to advance evidence-building efforts, improve performance, and enable progress in meeting agencies' strategic objectives and outcomes.

The Federal Information Security Management (E-Government) Act of 2002, Public Law 107-347, 116 Stat 2899, December 17, 2002.

*How applicable*: The E-Government Act enhanced the management and promotion of electronic Government services and processes by establishing a Federal Chief Information Officer within the Office of Management and Budget and by further establishing, under 44 U.S.C. 3506, Federal agency CIO responsibilities for information resources management planning, open data planning, data asset usage, data information inventories, public data assets, and a "broad framework of measures that require using Internet-based information technology to enhance citizen access to Government information and services, and for other purposes."

*Why important*: The E-Government Act established Federal agencies' responsibility for ensuring agency compliance with and prompt, efficient, and effective implementation of the information policies and information resources management responsibilities established under this subchapter, including the reduction of information collection burdens on the public.