



**Congressional
Research Service**

Informing the legislative debate since 1914

Federal Data Management: Issues and Challenges in the Use of Data Standards

April 29, 2024

Congressional Research Service

<https://crsreports.congress.gov>

R48053



R48053

April 29, 2024

Natalie R. Ortiz
Analyst in Government
Organization and
Management

Federal Data Management: Issues and Challenges in the Use of Data Standards

The federal government manages a significant amount of data. Congress expects federal agencies to be effective managers and stewards of the data with which they are entrusted. Sometimes, Congress contributes to this stewardship by enacting policies that require data standards. Data standards can be found in laws concerning a wide array of congressional matters and interests. When Congress requires data standards, it is essentially requiring some degree of data governance and management. In the most general sense, data standards create rules for data in some way and, when used, could contribute to the usability of data by agencies, including to their consistency, transparency, discoverability, reliability, accessibility, and quality. Data standards may also reduce administrative burdens, assist in extracting value from the federal investment in information technology (IT), support federal data integration and interoperability, and contribute to the use of emerging technologies, such as artificial intelligence.

Data standards are numerous and varied, and data can be subjected to standards in many ways. Various types of data standards appear in the law, such as open data standards, data format standards, and data exchange standards. Some of these types can be interpreted in practice to have multiple meanings, and some standardize data in some ways but not necessarily in others, which can affect the extent to which they serve Congress's intended policy goals. There are also differences among agencies in how they define and describe what data standards do independent of a legislative mandate to use them. Put simply, there is no single way to approach data standards, which can complicate their effective use. Thus, specifying data standards in the law can pose a challenge to lawmakers. Within agencies, chief data officers (CDOs) have a statutory responsibility for data management. If Congress identifies a role for data standards in future legislation—whether for specific programs, regulatory matters, or government-wide operations—it may consider ensuring that CDOs have a role in implementing these data standards and are sufficiently resourced to manage the use of data standards within agencies and the amount of data that agencies oversee.

Data standards operate within separate policy frameworks for information resource management and technical standards. Technical standards establish specifications for products or production methods, and data standards for federal purposes have largely been viewed as being under the umbrella of technical standards. Federal policymakers have generally preferred to use voluntary consensus standards—that is, technical standards developed by the private sector—in lieu of standards developed by the federal government for its own unique purposes. Voluntary consensus standards are not always available and, in some cases, may fail to serve agencies' data-related needs. While agencies receive some guidance on implementing data standards from the Office of Management and Budget and the National Institute of Standards and Technology, this guidance and the underlying policy frameworks may require further coordination to help promote the effective management and use of data standards by agencies.

Congress has at times sought to enable data interoperability by enacting laws that require data standards. Some of the greatest challenges to achieving interoperability concern the coordination of people, organizations, and processes. Relying on data standards alone will probably be insufficient to address some of these barriers to data interoperability. Congress may consider including frameworks for data governance in policies where data interoperability is a goal.

Contents

Introduction	1
Using Data Standards: Benefits and Challenges	2
Benefits	2
Issues in Defining Data Standards	5
Inadequate Data Standards Management and Data Governance.....	6
Data Standards in Context: Frameworks and Types.....	7
Information Management.....	7
Data Management.....	8
Technical Standards.....	9
Voluntary Consensus Standards.....	10
Types of Data Standards in Federal Laws	12
Open Data Standards.....	13
Data Format Standards.....	14
Data Exchange Standards	14
Data Element Standards.....	16
Metadata and Metadata Standards	18
Implementing Data Standards for Federal Data	19
Office of Management and Budget	20
Data Standards for Information Resource Management.....	20
Data Standards for Federal Statistics	21
Data Standards and a Federal Data Strategy.....	22
National Institute of Standards and Technology	23
Information Technology Guidance	23
Recent Developments	25
Considerations for Congress.....	25
Policy Options for Data Governance and Data Management	26
Designating a Role for Chief Data Officers.....	26
Government-Wide Policy Options for Managing Data Standards	26
Policy Coordination	27
Designating a Role for NIST	28
Policy Options for Specifying Data Standards in the Law.....	28
Implications for Data Interoperability.....	28

Contacts

Author Information.....	30
-------------------------	----

Introduction

The federal government might be one of the largest producers of data in the world.¹ Congress entrusts federal agencies to be effective stewards of these data. Congress has a broad and long-standing interest not only in how agencies use data in day-to-day operations but in ensuring the judicious management of federal data. Data standards may enable federal data to be managed more effectively. When federal data are subject to data standards, there is an expectation for those data to conform to or comply with the standards' requirements. Conformity to those requirements should then send a signal about the readiness and usability of the data for a particular purpose.

For this report, *federal data* means data collected, processed, maintained, disseminated, managed, or regulated by a federal agency, including that which are reported to a federal agency. The Office of Management and Budget (OMB) has characterized the influence of federal data in the following way:

Federal data drives the U.S. economy and civic engagement and there is virtually no policy or program decision facing a federal agency that would not benefit from the use of data.²

Consistent with this influence, the public also expects agencies to effectively produce federal data. For example, the Commerce Department states that more than 30 million U.S. businesses, 325 million Americans, and 93,000 tribal, state, and local governments rely on its data to make informed decisions.³

Federal lawmakers sometimes require data standards for the data that agencies are entrusted to manage. Requirements for data standards can be found in laws concerning a variety of congressional matters and interests, including federal spending, financial regulation, homeland security, public land use, transportation, environmental protection and conservation, public welfare, health care, and others. While data standards generally establish rules for data, there are various data standards and various ways data could be subjected to standards. What exactly data standards require for data may largely depend on context and implementation, which includes how data standards have been specified by lawmakers within any given law.

Data standards can be used for an array of purposes, factoring into how federal agencies use data in various operations and processes, including in the management of various federal programs. For example, these standards might set requirements for what data are needed at a minimum to automate a payment process,⁴ what “primary place of performance” is to mean for all agencies when reporting on federal financial awards,⁵ what data certain users of an information system can access, or how data are to be structured or formatted to electronically transmit data across information systems using a particular machine-readable format.

This report provides an overview of data standards for federal data, focusing primarily on their relationship to data management. It discusses a number of issues in using data standards,

¹ Bill Brantley, “The Value of Federal Data,” March 18, 2018, <https://digital.gov/2018/03/14/data-briefing-value-federal-government-data/>.

² OMB, *M-19-18: Federal Data Strategy—A Framework for Consistency*, June 4, 2019, p. 1, <https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf>.

³ Department of Commerce, “America’s Data Agency,” <https://www.commerce.gov/tags/americas-data-agency>.

⁴ This example draws from the Federal Integrated Business Framework. See more at General Services Administration (GSA), “Standard Data Elements—Financial Management,” https://ussm.gsa.gov/fibf-fm/#standard_data_elements.

⁵ This example is from the implementation of the data standards required by the Digital Accountability and Transparency Act (P.L. 113-101). For additional context, see Government Accountability Office, “DATA Act,” <https://www.gao.gov/assets/680/674866.pdf>.

including how to define *data standards* and the consequences of inadequate data standards management on the usability of federal data. Various types of data standards appear in the law—such as open data standards and data exchange standards—but such types can be loose constructions in practice and interpreted in various ways. Some of the ways data standards have been specified, implemented, and used in federal policymaking activities are discussed below.

This report also discusses the multiple policy frameworks for data standards. OMB and the National Institute of Standards and Technology (NIST) provide guidance to agencies on using data standards within these frameworks. The report concludes with some considerations for Congress when it seeks to require data standards for federal data, including the specification of data standards, data standards within specific policies, data standards for federal data government-wide, and federal data interoperability.⁶

Using Data Standards: Benefits and Challenges

Congress may include data standards in laws to ensure that federal data are useful. Independent of a legislative mandate, agencies may also use data standards to increase the usefulness of the federal data they manage. Where there is a desire within Congress to use data to inform its own efforts to develop policy, data standards may have a role in ensuring that data from executive branch agencies are useful for the lawmaking process. This section begins with a discussion of data standards as a data governance activity and some of the ways that data standards contribute to the usefulness of federal data according to these legislative interests and administrative efforts. Following a discussion of the benefits of using data standards, some of the challenges to using them are described.

Benefits

When policymakers create requirements for data standards in a law, they are expecting some degree of data governance and management. In a report that stems from several statutory requirements for it to report on the quality of certain federal data and on certain efforts related to federal data management, the Government Accountability Office (GAO) states that an effective data governance framework for federal data allows agencies to improve performance and outcomes.⁷ Data governance includes the authorities, roles, responsibilities, organizational structures, policies, procedures, standards, and resources for the definition, stewardship, production, security, and use of data.⁸ GAO suggests that data governance is distinct from but a precursor to effective data management, because the former is concerned with establishing mechanisms for decisionmaking while the latter is concerned with the implementation of those decisions.⁹

⁶ In the context of information technology, *interoperability* generally refers to the ability of data from one system to be used by another system. The emphasis in interoperability is on the extent to which data is ready for use following data exchange. Some laws have defined *interoperability* for a specific purpose (e.g., E-Government Act of 2002 [P.L. 107-347]), while other laws use the term without a definition (e.g., Paperwork Reduction Act of 1995 [P.L. 104-13]).

⁷ GAO, *Data Governance: Agencies Made Progress in Establishing Governance, but Need to Address Key Milestones*, GAO-21-152, December 2020, p. 7, <https://www.gao.gov/assets/gao-21-152.pdf#page=12>.

⁸ GAO, *Data Governance*, pp. 4-6.

⁹ In a somewhat different construction, GSA suggests that data governance is the practice of data management (see GSA, “Ten Plays of Our Data and Analytics Approach,” October 2020, <https://coe.gsa.gov/2020/10/19/da-update-9.html#Key%20Concepts>). One industry-based group describes data governance as the exercise of authority and control over the management of data assets and states that all organizations make such decisions regardless of whether (continued...)

GAO states that data governance ensures that data are transparent, accessible, and of sufficient quality for their intended use.¹⁰ In a separate, congressionally mandated report, GAO identifies data standards as “a recognized approach for increasing the consistency, and therefore the transparency, of data,” contributing to completeness, accuracy, and usefulness.¹¹ GAO also reported that data standards could reduce burden by minimizing inconsistent and duplicative reporting requirements. While Congress has a general interest in reducing administrative burdens on the public, including paperwork burdens, data standards may result in efficiencies for agencies and allow agencies to meaningfully aggregate or compare data that is otherwise difficult in the absence of data standards.

Congress sometimes requires data standards for certain government-wide activities and operations, in part to promote the transparency of federal data. The Digital Accountability and Transparency Act of 2014 (DATA Act), for example, required data standards for the financial data reported by executive branch agencies, in part to produce reliable government-wide spending data on USAspending.gov.¹² Title II of the Foundations for Evidence-Based Policymaking Act of 2018 (FEBPA)—known as the Open, Public, Electronic, and Necessary (OPEN) Government Data Act—requires OMB to develop guidance for using metadata to describe agency datasets in comprehensive data inventories.¹³ This guidance could include using metadata standards that are consistent with administrative practices that the act sought to codify.¹⁴ Under the OPEN Government Data Act, agencies must make certain agency datasets and the corresponding metadata available to the public.¹⁵ In addition to making federal data discoverable to the public, the transparency of datasets maintained by a federal agency may also assist other federal agencies

data governance is a formalized function (see DAMA International, *DAMA-DMBOK: Data Management Book of Knowledge*, 2nd ed. [Basking Ridge, NJ: Technics Publications, 2017], p. 67). GAO’s definition of *data governance* may suggest more formalized functions.

¹⁰ GAO, *Data Governance*, p. 4.

¹¹ GAO, *Grants Management: Action Needed to Ensure Consistency and Usefulness of New Data Standards*, GAO-24-106164, January 2024, pp. 4, <https://www.gao.gov/assets/d24106164.pdf#page=9>.

¹² One of the purposes of the DATA Act is to “establish Government-wide data standards for financial data and provide consistent, reliable, and searchable Government-wide spending data that is displayed accurately for taxpayers and policy makers on USAspending.gov (or a successor system that displays the data)” (P.L. 113-101, §2(2); 128 Stat. 1146).

¹³ P.L. 115-435, §202(d); 132 Stat. 5538. *Comprehensive* in this context includes datasets created by, collected by, under the control of, at the direction of, or maintained by an agency. The inventory is supposed to provide a clear and comprehensive understanding of the datasets in the possession of an agency (see 44 U.S.C. §3511(a)(1-2)). The act defines and uses the term *data assets* to mean “a collection of data elements or data sets that may be grouped together” (44 U.S.C. §3502(17)).

¹⁴ This codification is discussed in U.S. Congress, House Committee on Oversight and Government Reform, *Foundations for Evidence-Based Policymaking Act of 2017*, report to accompany H.R. 4174, H.Rept. 115-411, 115th Cong., 1st sess., pp. 11-12, <https://www.congress.gov/115/crpt/hrpt411/CRPT-115hrpt411.pdf#page=11>. The committee report discusses Executive Order 13642, “Making Open and Machine Readable the New Default for Government Information” (available at <https://www.govinfo.gov/content/pkg/FR-2013-05-14/pdf/2013-11533.pdf>) and at OMB, *M-13-13: Open Data Policy—Managing Information as an Asset*, May 9, 2013, https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2013/m-13-13.pdf. This memorandum introduced the use of metadata standards. OMB identifies these standards by the name “Project Open Data Metadata Schema.” Information about these standards (schema) can be found at <https://resources.data.gov/resources/dcat-us/> and are also discussed elsewhere in this report.

¹⁵ P.L. 115-435, §202(d); 132 Stat. 5539; 44 U.S.C. §3511(a)(2)(C-D). For additional information on the OPEN Government Data Act, see CRS In Focus IF12299, *The OPEN Government Data Act: A Primer*, by Meghan M. Stuessy.

in determining whether collecting data from the public or purchasing data may duplicate data available elsewhere within the federal government.¹⁶

Congress and the policymaking process may also benefit from data standards. In the 118th Congress, H.Con.Res. 49 would establish a commission within the legislative branch that would study and recommend how Congress might use “real-time, structured, integrated, and machine-readable” data in lawmaking.¹⁷ Use of data in this way for Congress’s needs may depend on a number of data management activities that occur within and across agencies, such as the effective use of data standards.

Administratively, some agencies have signaled their plans to use data standards in their agency-wide, or enterprise, data strategies.¹⁸ For example, the Department of State stated in its 2021 data strategy that it would transform how it collectively manages and uses data across its mission areas. The department identified defining and implementing data standards as part of a goal to establish mission-driven data management, saying, “An enterprise approach to data standards is needed, as current approaches are bespoke to specific data products and are not applied uniformly nor broadly understood. . . . The standards will enable greater discovery, utility, security, and efficacy of the department’s data.”¹⁹

Data standards for federal data are also relevant to emerging technologies affecting federal government operations, particularly artificial intelligence (AI). NIST characterized data standards as making the training data needed for machine learning applications more visible to users and more usable, assisting in creating “effective, reliable, robust, and trustworthy AI technologies.”²⁰ In this case, data standards are viewed as “measuring and sharing information relating to the quality, utility, and access of datasets . . . assist[ing] potential users in making informed decisions about the data’s applicability to their purpose, and help[ing] prevent misuse.” OMB’s guidance to agencies indicates that any data used to develop, test, or maintain AI applications should be assessed for quality and other features and that reducing barriers to data for training, testing, and operating AI should be supported by resources that enable data governance and management practices, particularly in data collection, curation, labeling, and stewardship.²¹

¹⁶ The Paperwork Reduction Act of 1995 (PRA; P.L. 104-13) requires agencies to complete certain steps before collecting information from the public, including receiving OMB’s approval of the information collection (44 U.S.C. §3507(a)). OMB requires agencies to provide certifications of certain criteria relevant to the information collection, including that it “is not unnecessarily duplicative of information otherwise reasonably accessible to the agency” (5 C.F.R. §1320.9(b)). For more information on the PRA’s information collection requirements, see CRS In Focus IF11837, *The Paperwork Reduction Act and Federal Collections of Information: A Brief Overview*, by Maeve P. Carey and Natalie R. Ortiz.

¹⁷ H.Con.Res. 49 §2(g)(1)(E).

¹⁸ For example, see National Aeronautics and Space Administration, *NASA Data Strategy*, January 2021, p. 18, https://www.nasa.gov/wp-content/uploads/2023/02/nasa_data_strategy.pdf#page=20; see also Department of Labor, *Enterprise Data Strategy*, April 2022, p. 8, <https://www.dol.gov/sites/dolgov/files/Data-Governance/DOL-Enterprise-Data-Strategy-2022.pdf#page=10>.

¹⁹ Department of State, *Enterprise Data Strategy: Empowering Data Informed Diplomacy*, September 2021, p. 14, <https://www.state.gov/wp-content/uploads/2021/09/Reference-EDS-Accessible.pdf#page=14>.

²⁰ NIST, *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, August 9, 2019, p. 13, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf#page=15.

²¹ OMB, *M-24-10: Advancing Governance, Innovation, and Risk Management for Agencies*, , March 28, 2024, p. 11, <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf#page=11>.

Issues in Defining Data Standards

OMB, the General Services Administration (GSA), and the National Archives and Records Administration (NARA)—which jointly maintain online tools, guidance, and other resources for managing and using federal data—have characterized the universe of data standards as “large, varied, and complex” and have stated there is no single definition for *data standards* for federal data management purposes.²²

Researchers have found that practitioners sometimes disagree about concepts fundamental to data standards.²³ Uncertainty in fundamentals might be a challenge when data standards often require coordination among different groups. Some have suggested that effectively implementing data standards requires different groups to understand data standards, find a shared vocabulary, and figure out collectively how to develop and use data standards that serve related but not identical goals.²⁴

Federal agencies differ in how they define data standards. For example, the Environmental Protection Agency defines *data standards* as “documented agreements on representation, format, definition, structuring, tagging, transmission, manipulation, use, and management of data.”²⁵ The U.S. Geological Survey states that data standards operate at the “parameter-level” and “dataset-level.”²⁶ As part of their collective effort, OMB, GSA, and NARA define *data standards* as a hierarchy of concepts while also noting the lack of formal definitions for the concepts within this hierarchy (see text box “Data Standards as a Hierarchy of Concepts”). Thus, this hierarchy also represents just one interpretation of data standards that may not be widely used outside of these agencies or are termed differently in practice.

Data Standards as a Hierarchy of Concepts

GSA, OMB, and NARA characterize data standards as often being comprised of “smaller component pieces,” “interchangeable parts,” or “common building blocks” that can be mixed and matched for different purposes. In this way, the agencies break down data standards into a “hierarchy of concepts” with each building on the previous:

- **Data standard components.** Data standards are typically made up of discrete “components.” Components of a data standard can include specification for (1) data type; (2) identifiers; (3) vocabulary, such as terms and definitions; (4) data models, schemas, and other representations that define relationships among pieces of information; (5) data format; and (6) protocols for reading and writing data within a file system or database or across computer systems and networks.
- **Data standards package.** Multiple components can be assembled together, creating a more comprehensive “data standards package.” These packages provide the instructions on how to implement the individual components. The agencies note that the exact terminology for “data standards package” varies.

²² OMB, GSA, and NARA, “Data Standards Concepts and Definitions,” <https://resources.data.gov/standards/concepts/>.

²³ Boris Otto, Erwin Folmer, and Verena Ebner, “A Characteristics Framework for Semantic Information Systems Standards,” *Information Systems and e-Business Management*, vol. 10, no. 4 (December 2012), p. 573.

²⁴ Richard Berner and Kathryn Judge, *The Data Standardization Challenge*, European Corporate Governance Institute, ECGI Working Paper Series in Law no. 438/2019, January 2019, p. 12.

²⁵ Environmental Protection Agency, “Learn About Data Standards,” <https://www.epa.gov/data-standards/learn-about-data-standards#What>.

²⁶ U.S. Geological Survey, “Data Standards,” <https://www.usgs.gov/data-management/data-standards>. The agency describes dataset-level standards as specifying the “scientific domain, structure, relationships, field labels, and parameter-level standards for the dataset as a whole.” These dataset-level standards are normally documented in a data dictionary. Parameter-level standards specify “the format and units for a given parameter or field within a dataset and help users correctly interpret the values” and “should be adopted at the time of data collection ... when values in a field are created or recorded.”

- **Data standards framework.** The agencies describe a data standards framework as a “comprehensive system of reusable data standards components” that allow individual components to be combined in various ways in order to serve a range of use cases.

GSA, OMB, and NARA note that their concepts do not represent formal definitions but are generally used to describe common concepts associated with data standards. As such, it may be difficult to directly apply these concepts and how they relate to each other in other contexts.

Additionally, the agencies note that *standard* itself has two meanings, one that is used in government to refer to a requirement, a compliance measure, or a minimum set of qualification criteria that something must meet and a second that is used in digital technology to refer to a common technical specification for how information is described, processed, or transmitted. For their purposes, GSA, OMB, and NARA use *standard* according to this second meaning, because “data standards are not intended to describe minimum qualification criteria that data should meet, but instead describe technical specifications that allow for the consistent and interoperable collection and exchange of data in specific environments.”²⁷

GAO differentiates between data standards that are used by programs or individual agencies and government-wide data standards.²⁸ In the former, an agency or program may use agreed-upon definitions and technical specifications of data elements that may be different from those used at another agency or for another program. For example, different agencies could use the same name for a data element (e.g., address), but that data element may actually represent different things at each agency (e.g., mailing address versus residential address). In contrast, government-wide data standards attempt to consistently define and specify data elements across agencies or programs.

Some laws require specific types of data standards (e.g., data exchange standards, data format standards, or open data standards). In practice, these types of data standards also do not have a single, universally applied definition. Different interpretations of data standards may lead to differences in their implementation. Some data standards that appear in the law and the various ways such standards can be interpreted and have been implemented are discussed in a later section of this report (see “Types of Data Standards in Federal Laws”).

Inadequate Data Standards Management and Data Governance

Data standards for federal data may require deliberate and active management to contribute to their intended purposes. Inadequate management of data standards has real consequences on the reliability, discoverability, and usability of federal data.

GAO has found shortcomings in the data that appear on USAspending.gov, which is supposed to be an authoritative source of federal spending information for the public and policymakers.²⁹ It found that the federal financial data standards developed by OMB and the Treasury Department pursuant to the DATA Act were not uniformly interpreted or applied by agencies, resulting in reporting differences that affect the reliability of information on USAspending.gov.³⁰ GAO, thus, recommended the consistent use of federal financial data standards to ensure the integrity of the standards over time. Specifically, GAO reported that a properly implemented formal data governance structure could help adjudicate revisions to the data standards, ensure compliance

²⁷ OMB, GSA, and NARA, “Data Standards Concepts and Definitions.”

²⁸ GAO, *Grants Management: Action Needed to Ensure Consistency and Usefulness of New Data Standards*, p. 4.

²⁹ GAO has published several reports on the data available through USAspending.gov. GAO provides a list of some of these reports, in addition to other relevant reports, in GAO, *Federal Spending Transparency: Opportunities Exist to Further Improve the Information Available on USAspending.gov*, GAO-22-104702, November 2021, pp. 59-61, <https://www.gao.gov/assets/d22104702.pdf#page=65>.

³⁰ GAO, *DATA Act: Quality of Data Submissions Has Improved but Further Action Is Needed to Disclose Known Data Limitations*, GAO-20-75, November 2019, p. 27, <https://www.gao.gov/assets/gao-20-75.pdf#page=33>.

with the data standards, and increase the accuracy of the federal financial data made available to the public.³¹

There is also some indication of inadequate management of metadata standards agencies use for making data available to the public. The Federal Chief Data Officers (CDO) Council recommended that OMB establish governance of the metadata standards it requires agencies to use to catalogue federal data pursuant to the OPEN Government Data Act.³² In 2013, OMB introduced metadata standards—which OMB called a schema—in association with an administrative effort to manage federal data that would support their discoverability and usability by the public.³³ The underlying metadata standards have since been updated by their nongovernmental developers, who stated that the update was necessary because “the original specification lacked a number of essential features.”³⁴ The council recommended that OMB adopt the updated metadata standards because they were more applicable to federal agencies and their constituents, improving dataset “search, discovery, and appropriate use.”³⁵ Additionally, some stakeholders believe that the current metadata standards for open federal data “will not capture the rich detail” needed for adequate transparency for federal statistical data products but are still a “useful starting point” for other federal data products, in part because most federal agencies have started to use the existing standards.³⁶

Data Standards in Context: Frameworks and Types

Data standards for federal data operate within policy frameworks for federal information management and the use of technical standards by agencies. Federal information management policies have evolved to specifically include data management and a role for CDOs to oversee data management within agencies. Within technical standards policy, there is preference for agencies to use voluntary consensus standards over other types of technical standards.

Information Management

One framework for using data standards includes the Paperwork Reduction Act (PRA)—codified at Title 44, Sections 3501-3521, of the *U.S. Code*—which governs the collection, processing, and management of federal data by agencies.³⁷ Among several other purposes, the PRA was designed to improve the quality and use of federal data and to minimize the federal cost of collecting, managing, and using federal data.³⁸ Through the amendments made by FEBPA to the *U.S. Code*

³¹ GAO, *DATA Act: OMB Needs to Formalize Data Governance for Reporting Federal Spending*, GAO-19-284, March 2018, p. 4, <https://www.gao.gov/assets/gao-19-284.pdf#page=8>.

³² Federal CDO Council Data Inventory Working Group, *Enterprise Data Inventories*, April 2022, p. 20, https://resources.data.gov/assets/documents/CDOC_Data_Inventory_Report_Final.pdf#page=24. Requirements for a federal data catalogue are specified in Title 44, Section 3511(c), of the *U.S. Code*.

³³ OMB, *M-13-13*, p. 5.

³⁴ Riccardo Albertoni et al., eds., “Data Catalog Vocabulary (DCAT)—Version 2,” World Wide Web Consortium, February 4, 2020, <https://www.w3.org/TR/vocab-dcat-2/#motivation>.

³⁵ Federal CDO Council Data Inventory Working Group, *Enterprise Data Inventories*, April 2022, p. 20, https://resources.data.gov/assets/documents/CDOC_Data_Inventory_Report_Final.pdf#page=24.

³⁶ National Academies of Sciences, Engineering, and Medicine, *Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies* (Washington, D.C.: National Academies Press, 2022), p. 63, available for download at <https://doi.org/10.17226/26360>.

³⁷ Originally enacted as the Paperwork Reduction Act of 1980 (P.L. 96-511), it was amended by the Paperwork Reduction Act of 1995 (P.L. 104-13).

³⁸ P.L. 104-13, §2; 109 Stat. 164; 44 U.S.C. §3501(4-5).

provisions codifying the PRA, Congress has set some expectations for federal data management and for the use of data standards for federal data.

The PRA establishes government-wide and agency roles and responsibilities for several different aspects of information resource management (IRM). The PRA defines IRM as the “process of managing information resources to accomplish agency missions and to improve agency performance, including through the reduction of information collection burdens on the public.”³⁹ Some have noted that while many observations of the PRA are primarily concerned with how it governs the collection of information from individuals, businesses, and other entities, the act’s scope is to improve management and efficiency in the federal government.⁴⁰

Data Management

Federal agencies’ use of data standards is closely linked with information technology (IT) and information systems. A congressional committee report on the PRA noted that one intent of the PRA’s IRM requirements was “to refocus the attention of federal managers on the pressing need to use [IT] to support programs efficiently and effectively:”

The reduction of public paperwork burdens will also be served by the legislation’s other management focus. The still widening gap between possibilities for improved government operations through the use of information technology, and the government’s apparent inability to take advantage of this technology, demonstrates that the [Paperwork Reduction Act of 1980]’s IRM mandates have not been sufficiently realized. Today’s information systems offer the government unprecedented opportunities to provide higher quality services tailored to the public’s changing needs, delivered more effectively, faster, at lower cost, and with reduced burdens on the public. Unfortunately, Federal agencies have not kept pace with evolving management practices and skills necessary to: (1) precisely define critical information needs; and (2) select, apply, and manage changing information technologies. The result, in many cases, has been wasted resources, a frustrated public unable to get quality service, and a government ill-prepared to measure and manage its affairs in an acceptable manner. Despite spending more than \$200 billion on information management and systems in the past 12 years, the government has too little evidence of meaningful returns. The consequences—poor service quality, high costs, low productivity, unnecessary risks and burdens, and unexploited opportunities for improvement—cannot be tolerated.⁴¹

Despite the link, the House Committee on Oversight and Government Reform distinguished between data management and IT management when discussing the roles of chief information officers and CDOs for the purposes of FEBPA and its amendments to the *U.S. Code* provisions that codify the PRA. The committee reported, “Data management, as opposed to IT management or IT security, is about establishing effective procedures, standards, and controls to ensure quality, accuracy, transparency, and privacy of data.”⁴²

³⁹ P.L. 104-13, §2; 109 Stat. 166; 44 U.S.C. §3502(7). For the purpose of the PRA, the term *information resources* is defined to mean information and related resources such as personnel, funds, and information technology (44 U.S.C. §3502(6)).

⁴⁰ Robert Gellman, “Crowdsourcing, Citizen Science, and the Law: Legal Issues Affecting Federal Agencies,” Woodrow Wilson International Center for Scholars, p. 27, https://www.wilsoncenter.org/sites/default/files/media/documents/publication/CS_Legal_Barriers_Gellman.pdf#page=28.

⁴¹ U.S. Congress, Senate Committee on Governmental Affairs, *Paperwork Reduction Act of 1995*, report to accompany S. 244, 104th Cong., 1st sess., S.Rept. 104-8, February 14, 1995, p. 5, <https://www.congress.gov/104/crpt/srpt8/CRPT-104srpt8.pdf#page=8>.

⁴² U.S. Congress, House Committee on Oversight and Government Reform, *Foundations for Evidence-Based* (continued...)

OMB is tasked with developing, coordinating, and overseeing government-wide IRM policies, principles, standards, and guidelines.⁴³ The role of OMB is discussed later in this report (see “Implementing Data Standards for Federal Data”). In addition, under specific laws, OMB has a role in establishing—or working with other agencies to establish—data standards in relation to specific federal government management issues.⁴⁴

Chief Data Officers

Title II of FEBPA—the OPEN Government Data Act—established the role of a CDO within each agency.⁴⁵ CDOs are responsible for data management.⁴⁶ The role centralizes data management within an agency. The House Committee on Oversight and Government Reform believed that the centralized management of data would improve interoperability and enhance the transparency of existing federal data.⁴⁷

Among other functions, an agency’s CDO is responsible for (1) supporting efforts within the agency to use data for performance improvement;⁴⁸ (2) ensuring that agency data conform, to the extent practicable, with data management best practices;⁴⁹ and (3) standardizing the data formats of federal datasets.⁵⁰ CDOs are also responsible for certain agency IRM activities specified in the PRA.⁵¹ This includes improving the utility of agency data to all users within and outside the agency and the efficient and effective management and use of the data that are collected by the agency from the public.⁵²

Technical Standards

Technical standards are performance-based or design-based specifications, including the practices to manage these specifications.⁵³ Technical standards establish, among other criteria, terminology, rules, and specifications for products and production methods.⁵⁴ In general, the policy framework for these standards is the National Technology Transfer and Advancement Act of 1995

Policymaking Act of 2017, report to accompany H.R. 4174, H.Rept. 115-411, 115th Cong., 1st sess., p. 16, <https://www.congress.gov/115/crpt/hrpt411/CRPT-115hrpt411.pdf#page=16>.

⁴³ P.L. 104-13, §2; 109 Stat. 167; 44 U.S.C. §3504(a)(1)(A).

⁴⁴ This includes the DATA Act (P.L. 113-101, §4(a)(1); 128 Stat. 1148) and the Grant Reporting Efficiency and Agreement Transparency Act (P.L. 116-103, §4; 133 Stat. 3268; 31 U.S.C. §6402(a)(2)), among others.

⁴⁵ P.L. 115-435, §202(e); 132 Stat. 5541; 44 U.S.C. §3520(a).

⁴⁶ 44 U.S.C. §3520(c)(1).

⁴⁷ U.S. Congress, House Committee on Oversight and Government Reform, *Foundations for Evidence-Based Policymaking Act of 2017*, report to accompany H.R. 4174, H.Rept. 115-411, 115th Cong., 1st sess., p. 16, <https://www.congress.gov/115/crpt/hrpt411/CRPT-115hrpt411.pdf#page=16>.

⁴⁸ 44 U.S.C. §3520(c)(8). This section makes a reference to Title 31, Section 1124(a)(2), of the *U.S. Code*, which describes the function of an agency’s performance improvement officer, of which the CDO is expected to support with data.

⁴⁹ 44 U.S.C. §3520(c)(6).

⁵⁰ 44 U.S.C. §3520(c)(3).

⁵¹ P.L. 115-435, §202(e); 132 Stat. 5541; 44 U.S.C. §3520(c)(5). This includes responsibility for activities described in Title 44, Sections 3506(b-d), 3506(f), 3506(i), and 3511, of the *U.S. Code*.

⁵² These requirements are described in Title 44, Sections 3506(b)(1)(C) and 3506(c)(1)(A)(vi), of the *U.S. Code*.

⁵³ P.L. 104-113, §12(d)(4); 110 Stat. 783.

⁵⁴ OMB, *Circular A-119: Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities*, January 27, 2016, p. 15, https://www.whitehouse.gov/wp-content/uploads/2020/07/revised_circular_a-119_as_of_1_22.pdf.

(NTTAA).⁵⁵ OMB provides further instructions to agencies on using technical standards in Circular A-119, *Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities*.

Data standards for federal data have been viewed as being under the umbrella of technical standards. This may be in part because the use of data standards for federal data largely emerged along with the federal government's use of computers in the 1960s and subsequent efforts to use technical standards for federal IT. OMB has associated data standards with the prevailing policy on technical standards for federal information processing and dissemination policies, including interoperability among information systems:

Consistent with existing policies relating to Federal agencies' use of standards [under *Circular A-119*] for information as it is collected or created, agencies must use standards in order to promote data interoperability and openness.⁵⁶

Voluntary Consensus Standards

The NTTAA and *Circular A-119* prefer technical standards that have been developed by standards development organizations (SDOs) in the private or nongovernment sectors.⁵⁷ Statute and policy-related documents might formally refer to these SDOs as *voluntary consensus standards bodies*, which OMB defines as entities that plan, develop, establish, or coordinate voluntary consensus standards using a specific development process (see the text box "Developing Voluntary Consensus Standards").⁵⁸ Sometimes, statute requires data standards to incorporate standards developed or maintained by voluntary consensus standards bodies.

NIST is generally involved in coordinating the use by agencies of technical standards developed by private sector SDOs.⁵⁹ NIST generally views technical standards as documentary standards, meaning they are published in documents that detail their requirements.⁶⁰ NIST's role in assisting agencies in using data standards is discussed in "Implementing Data Standards for Federal Data."

Voluntary consensus standards are intended to:

- eliminate costs to the federal government to develop its own standards, decreasing procurement costs and the burden to comply with agency regulations;
- incentivize standards for needs that are national in scope, encouraging long-term growth for U.S. enterprises and promoting efficiency and competition; and

⁵⁵ P.L. 104-113, §12(a)(3); 110 Stat. 782.

⁵⁶ OMB, *M-13-13*, p. 7.

⁵⁷ OMB, *Circular A-119*, pp. 17-19.

⁵⁸ The NTTAA directed all federal agencies to use technical standards developed or adopted by *voluntary consensus standards bodies* but did not formally define them (P.L. 104-113, §12(d)(1); 110 Stat. 783). *Circular A-119* defines a voluntary consensus standards body as "a type of association, organization, or technical society that plans, develops, establishes, or coordinates voluntary consensus standards using a voluntary consensus standards development process" (p. 16).

⁵⁹ P.L. 104-113, §12(a)(3); 110 Stat. 782; 15 U.S.C. §272(b)(3). The NTTAA added Section 272(b)(3) to Title 15 of the *U.S. Code*. Language in that section was amended by Section 403(2) of P.L. 114-329 (130 Stat. 3023), but the amended language did not affect the function of NIST to "coordinate the use by Federal agencies of private sector standards, emphasizing where possible the use of standards developed by private, consensus organizations."

⁶⁰ NIST, "Documentary Standards," <https://www.nist.gov/feature-stories/why-you-need-standards/documentary-standards>.

- further government reliance on the private sector’s expertise to supply cost-efficient products and services.⁶¹

The American National Standards Institute—a private nonprofit organization that the executive branch has recognized as the coordinator of the U.S. standardization system⁶²—formally accredits U.S.-based SDOs.⁶³ Standards developed by these accredited organizations are promoted as demonstrating compliance with the definition of *voluntary consensus standard* in accordance with the NTTAA and *Circular A-119*.⁶⁴

Developing Voluntary Consensus Standards

Circular A-119 describes voluntary consensus standards bodies as using a development process with certain attributes:

- (i) Openness: The procedures or processes used are open to interested parties. Such parties are provided meaningful opportunities to participate in standards development on a non-discriminatory basis. The procedures or processes for participating in standards development and for developing the standard are transparent.
- (ii) Balance: The standards development process should be balanced. Specifically, there should be meaningful involvement from a broad range of parties, with no single interest dominating the decision-making.
- (iii) Due process: Due process shall include documented and publically available policies and procedures, adequate notice of meetings and standards development, sufficient time to review drafts and prepare views and objections, access to views and objections of other participants, and a fair and impartial process for resolving conflicting views.
- (iv) Appeals process: An appeals process shall be available for the impartial handling of procedural appeals.
- (v) Consensus: Consensus is defined as general agreement, but not necessarily unanimity. During the development of consensus, comments and objections are considered using fair, impartial, open, and transparent processes.⁶⁵

The NTTAA requires federal agencies to participate in the development of voluntary consensus standards by SDOs when it is in the public’s interest and compatible with an agency’s mission, authorities, priorities, and resources.⁶⁶ *Circular A-119* provides additional direction to agencies, including on informing the public of ongoing or planned participation.⁶⁷ OMB in 2012 emphasized private sector leadership in developing standards, noting that in some circumstances a federal agency may need to actively engage in standards development to accelerate technological advances and technology adoption, particularly when a substantial government investment is being made.⁶⁸

⁶¹ OMB, *Circular A-119*, p. 14.

⁶² The White House, *United States Government National Standards Strategy for Critical and Emerging Technology*, May 2023, p. 4, <https://www.whitehouse.gov/wp-content/uploads/2023/05/US-Gov-National-Standards-Strategy-2023.pdf>.

⁶³ American National Standards Institute, “ANSI Frequently Asked Questions,” <https://ansi.org/standards-faqs#ansi>.

⁶⁴ American National Standards Institute, “American National Standards (ANS) Introduction,” <https://ansi.org/american-national-standards/ans-introduction/overview#introduction>.

⁶⁵ OMB, *Circular A-119*, p. 16.

⁶⁶ P.L. 104-113, §12(d)(2); 110 Stat. 783.

⁶⁷ OMB, *Circular A-119*, p. 29.

⁶⁸ OMB, *M-12-08: Principles for Federal Engagement in Standards Activities to Address National Priorities*, January 17, 2012, p. 2, https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2012/m-12-08_1.pdf#page=2.

OMB specifically prefers agencies to use voluntary consensus standards rather than government-unique standards.⁶⁹ *Government-unique standards* are developed by the federal government specifically for its use—including in its regulations, procurements, or program areas—and are not generally used by the private sector unless required by a federal regulation, for federal procurement reasons, or in connection with participation in a federal program.⁷⁰

Exceptions to Use

While use of voluntary consensus standards is preferred, it is not mandatory.⁷¹ The NTTAA permits exceptions when it is inconsistent with applicable law or otherwise impractical.⁷² OMB states that there is flexibility to allow for agencies to best meet their missions. In addition, there may be no suitable voluntary consensus standards available for an agency to use. In such situations, OMB permits an agency to:

- use other suitable standards that “deliver favorable technical and economic outcomes (such as improved interoperability) and are widely used in the marketplace;”
- develop its own standards (i.e., government-unique standards);
- use already developed government-unique standards;
- solicit interest from qualified SDOs to develop standards; or
- develop standards using the processes of voluntary consensus standard bodies.⁷³

If an agency elects to use or develop government-unique standards, then it must transmit the reasons for using such standards to OMB through NIST.⁷⁴

Types of Data Standards in Federal Laws

Some laws specify certain types of data standards, such as (1) open data standards, (2) data format standards, (3) data exchange standards, (4) data element standards, and (5) metadata standards. The terminology for these data standards is not without ambiguity. Some have

⁶⁹ OMB, *Circular A-119*, p. 17.

⁷⁰ OMB, *Circular A-119*, p. 16.

⁷¹ OMB defines *use* to mean “incorporation of a standard in whole, in part, or by reference for procurement purposes; inclusion of a standard in whole, in part, or by reference in regulation(s); or inclusion of a standard in whole, in part, or by reference in other mission-related activities” (*Circular A-119*, p. 20).

⁷² P.L. 104-113, §12(d)(3); 110 Stat. 783. OMB defines *impractical* to include circumstances in which such use would fail to serve the agency’s regulatory, procurement, or program needs; be infeasible; be inadequate, ineffectual, inefficient, or inconsistent with the agency mission or the goals of using voluntary consensus standards; be inconsistent with a provision of law; or impose more burdens or be less useful than the use of another standard (*Circular A-119*, p. 20).

⁷³ OMB, *Circular A-119*, pp. 19-20.

⁷⁴ The NTTAA requires reporting from agencies to OMB on the use of technical standards not developed or adopted by voluntary consensus standards bodies (P.L. 104-113, §12(d)(3); 110 Stat. 783). Following enactment of the NTTAA, OMB issued a revised *Circular A-119* on February 19, 1998. The circular at the time established that agencies would submit their reports and reasons for using *government-unique standards*—defined in the circular at the time as developed by the government for its own uses—to OMB through NIST (see OMB, “OMB Circular A-119; Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities—Final Revision of Circular A-119,” February 19, 1998, 63 *Federal Register* 8557, <https://www.govinfo.gov/content/pkg/FR-1998-02-19/pdf/98-4177.pdf>). The 2016 revision to *Circular A-119* maintains this reporting structure among agencies, NIST, and OMB (see pp. 33-34). These reports can be found at <https://www.nist.gov/standardsgov/nttaa-agency-reports>.

characterized it as confusing and unclear and have also observed that different terms could be used for the same type of data standard.⁷⁵

In practice, federal data may be standardized in some ways but not necessarily in others. For example, requiring that data be formatted using XML does not also guarantee that the data's potential users will find it equally useful.⁷⁶ Similarly, some data standard types are not necessarily mutually exclusive. For example, XML is a data format standard and a data exchange standard; it is also an open standard. A metadata standard, which could also be called a schema, may define data elements in ways that are expected by an interpretation of data element standards.

Open Data Standards

Some data standards are *open*, meaning nonproprietary and available for use without a dependency on certain technologies or software applications.⁷⁷ *Open* may also refer to the ability to participate in standards development.⁷⁸ The concept can be interpreted in other ways.⁷⁹ Many data standards that are described as open facilitate the use of data in various ways, such as to establish a common language for data (e.g., definitions, identifiers, and code sets), the exchange of data (e.g., data formats such as XML), and data catalogs (e.g., online portals to publish datasets).⁸⁰ The text box “Examples of Open Data Standards in Federal Policy” notes some of the ways open data standards appear in federal law.

Examples of Open Data Standards in Federal Policy

Some laws specify the use of open data standards:

- The Bipartisan Budget Act of 2018 required data standards for federal reporting and state data exchanges that, to the extent practicable, incorporate “existing nonproprietary standards,” such as XML.⁸¹ XML is considered an “open standard” by its private sector developers, in part because (1) it was developed using an open process, and (2) it does not require any specific technologies or software to use.⁸²
- The OPEN Government Data Act defined *open government data asset* to mean, among other characteristics, a public dataset in an open format and based on an underlying open standard maintained by a standards organization.⁸³ In this context, being “based on an underlying open standard” could depend on what *open standard* means in a particular context, including any metadata standards that might describe such datasets, whereas *open format* could mean an open data format standard, such as XML. Data format standards and metadata standards are discussed below.

⁷⁵ Fenareti Lampathaki et al., “Business to Business Interoperability: A Current Review of XML Data Integration Standards,” *Computer Standards and Interfaces*, vol. 31 (2009), p. 1047.

⁷⁶ Lampathaki et al., “Business to Business Interoperability,” p. 1046.

⁷⁷ Andy Gower, “Open Standards vs. Open Source: A Basic Explanation,” IBM, April 2, 2019, <https://www.ibm.com/blog/open-standards-vs-open-source-explanation/>.

⁷⁸ Open Data Institute, “What Are Open Standards for Data?,” <https://standards.theodi.org/introduction/what-are-open-standards-for-data/>. See also the text box in this report “Developing Voluntary Consensus Standards” for a description of “openness” in this development process.

⁷⁹ Ken Krechmer, “Open Standards: A Call for Change,” *IEEE Communications Magazine*, vol 47, no. 5 (May 2009), p. 89, <https://ieeexplore.ieee.org/abstract/document/4939282>.

⁸⁰ Open Data Institute, “What Are Open Standards for Data?”

⁸¹ P.L. 115-123, §50606(a); 132 Stat. 230; 42 U.S.C. §711(h)(5)(B)(iii).

⁸² Erik T. Ray, *Learning XML* (Sebastopol, CA: O’Reilly and Associates, 2003), p. 8.

⁸³ P.L. 115-435, §202(a)(3); 132 Stat. 5535; 35 U.S.C. §3502(20).

Data Format Standards

Data format standards may refer to the syntax, encoding, and other often technical specifications that allow data to be stored, transferred, and then “read” or interpreted by a machine (e.g., a computer).⁸⁴ These types of standards include, for example, XML, comma separate values (CSV), and JavaScript Object Notation (JSON).⁸⁵ Some types of data formats may be recognizable as file types, but there are distinctions between data formats and file formats. Given that data format standards can structure data for transmission, some data format standards are efficient for data interoperability. The definition of *technical standards* in OMB’s *Circular A-119* includes formats for the exchange of information.⁸⁶

Machine-Readable in the Federal Context: Examples of Data Format Standards

- Data format standards are related to the concept of being machine-readable.⁸⁷ Federal datasets that can be made publicly available are expected, among other features, to be machine-readable.⁸⁸ *Machine-readable* was defined by the OPEN Government Data Act to mean “data in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost.”⁸⁹ Agency CDOs are responsible for standardizing data formats within their respective agencies.⁹⁰
- The House Committee on Oversight and Government Reform heard testimony that data format issues added to the costs of projects using federal data: “Most of the expense of big data projects comes from extracting information from different sources, transforming those data sets into the same format, and then loading them into new systems to be analyzed. If Federal data sets were consistently available using machine-readable formats to begin with, those expensive one-off projects would not be necessary.... When the government publishes its information, it needs to use non-proprietary data formats, formats that nobody owns.”⁹¹
- The U.S. Securities and Exchange Commission (SEC) reported that some machine-readable data is required by 38 out of the 52 statutorily required disclosures it oversees (i.e., forms, schedules, and statements). In 2009, the SEC required certain information to be provided in a machine-readable format and has continued to require structured data for various information collections. The SEC reports to Congress semiannually on public and internal uses of machine-readable data for corporate disclosures, including the costs and benefits of using machine-readable data.⁹²

Data Exchange Standards

Data exchange standards are generally conceived of as standards for transmitting (sending) and exchanging (sending and receiving) data across usually disparate systems. XML and JSON are two commonly used standards for exchanging data, particularly for those exchanges that occur over the internet.

⁸⁴ OMB, GSA, and NARA, “Data Standards Concepts and Definitions.”

⁸⁵ OMB, GSA, and NARA, “Data Standards Concepts and Definitions.”

⁸⁶ OMB, *Circular A-119*, p. 15.

⁸⁷ World Bank, “Open Data Essentials,” <https://opendatatoolkit.worldbank.org/en/data/opendatatoolkit/essentials#open-data>.

⁸⁸ P.L. 115-435, §202(a); 132 Stat. 5535; 44 U.S.C. §3502(20)(a).

⁸⁹ P.L. 115-435, §202(a); 132 Stat. 5535; 44 U.S.C. §3502(18).

⁹⁰ P.L. 115-435, §202(e); 132 Stat. 5541; 44 U.S.C. §3520(c)(3).

⁹¹ Testimony of Hudson Hollister, in U.S. Congress, House Committee on Oversight and Government Reform, *Legislative Proposals for Fostering Transparency*, hearing, 115th Cong., 1st sess., March 23, 2017, Serial No. 115-20, p. 7, <https://www.govinfo.gov/content/pkg/CHRG-115hhrg26499/pdf/CHRG-115hhrg26499.pdf>.

⁹² SEC, *Semi-Annual Report to Congress Regarding Public and Internal Use of Machine-Readable Data for Corporate Disclosures*, June 2023, pp. 1-2, <https://www.sec.gov/files/2023-fdta-report.pdf#page=5>.

Data exchange standards can facilitate data interoperability. For general IT purposes, *interoperability* refers to the ability of data from one system to be used by another system.⁹³ The text box below discusses one example of how data exchange standards specified in a law have been implemented. While data exchange and interoperability standards have generally been identified as a necessary component for sharing data in a generalized way, the specifics of these standards can vary considerably from case to case. As the text box illustrates, some data exchange standards establish requirements for data format (e.g., XML), while others may specify content, structure, and the meaning of data, suggesting a broad range of ways data exchange standards can present themselves.

Federal Data Exchange Standards: An Example of Implementation

In a section titled, “Data Exchange Standardization for Improved Interoperability,” the Middle Class Tax Relief and Job Creation Act of 2012 required the Department of Labor to establish data exchange standards for categories of information required under federal law for the unemployment compensation program.⁹⁴ The act requires the data exchange standards—to the extent possible—to be interoperable and nonproprietary, incorporating existing standards such as XML, and interoperable standards developed and maintained by:

- an international SDO,
- intergovernmental partnerships, such as the National Information Exchange Model (NIEM), and
- federal entities with certain financial-related authorities.⁹⁵

The department’s final rule designated XML as the data exchange standard, because it fulfilled many of the law’s requirements, but noted that interoperable standards developed and maintained by federal entities were not applicable to unemployment insurance processes.⁹⁶

The NIEM standards referenced in the act also rely on XML but also establish specifications for what NIEM calls “information exchange package documentation,” which defines the content, structure, and meaning of an information exchange message and is considered by NIEM as “the point of interoperability.”⁹⁷

NIEM standards may be more comprehensive in their design, enabling interoperability in a way that more narrowly specified data exchange standards cannot. In a joint effort, OMB, GSA, and NARA have referred to NIEM as an example of a “data standards package” and also a “data standards framework” within their hierarchy of data standards concepts, as discussed in the earlier text box “Data Standards as a Hierarchy of Concepts.”⁹⁸

There may be differences in what is intended by interoperability in laws where nonfederal operations are concerned (e.g., for state-administered federal programs) versus interoperability in federal interagency operations. The E-Government Act of 2002 defined *interoperability* to mean “the ability of different operating and software systems, applications, and services to communicate and exchange data in an accurate, effective, and consistent manner.”⁹⁹ The act applied the term for the purposes of pilot projects that would encourage federal data integration and data management to reduce the federal collection of duplicate data from the public, facilitate public access to federal data, and develop software applications that would reduce errors in

⁹³ For example, see the discussion in John Palfrey and Urs Gasser, *Interop: The Promise and Perils of Highly Interconnected Systems* (New York: Basic Books, 2012), pp. 5-7.

⁹⁴ P.L. 112-96, §2104; 126 Stat. 161; 42 U.S.C. §1111. For more on the program, see CRS In Focus IF10336, *The Fundamentals of Unemployment Compensation*, by Julie M. Whittaker and Katelin P. Isaacs.

⁹⁵ 42 U.S.C. §1111(a)(3).

⁹⁶ Department of Labor, Employment and Training Administration, “Federal-State Unemployment Insurance (UI) Program; Data Exchange Standardization as Required by Section 2104 of the Middle Class Tax Relief and Job Creation Act of 2012,” 79 *Federal Register* 9405, February 19, 2014, <https://www.govinfo.gov/content/pkg/FR-2014-02-19/pdf/2014-03496.pdf#page=2>. This final rule adopting XML is codified in Title 20, Sections 619.2(a) and 619.3(a), of the *Code of Federal Regulations*.

⁹⁷ NIEM, “IEPDs,” <https://niem.github.io/reference/iepd/>.

⁹⁸ OMB, GSA, and NARA, “Data Standards Concepts and Definitions.”

⁹⁹ P.L. 107-347, §101(a); 116 Stat. 2902; 44 U.S.C. §3601(6).

information electronically submitted to agencies.¹⁰⁰ One of the purposes of the act was to improve electronic and internet-based public services by improving the effectiveness and efficiency of interagency electronic processes and integrating related agency functions.¹⁰¹ The act directed OMB to develop a policy framework for IT standards that included the interoperability standards required by the PRA.¹⁰² While the PRA did not define *interoperability*, it directed OMB to promote federal data sharing through interoperability standards.¹⁰³ One way OMB has implemented the E-Government Act’s interoperability mandates is through the federal enterprise architecture (FEA), which is described as a framework and has evolved over time.¹⁰⁴ Among other features and components, the FEA provides a reference model for sharing data across the federal government and standardizing data exchanges.¹⁰⁵ In 2004, GAO testified that “hard work lies ahead to clarify and evolve the FEA and to ensure that well-managed architecture programs are actually established and executed, underscored, across the Government.”¹⁰⁶

Data Element Standards

Numerous statutes reference “data elements.”¹⁰⁷ Terms such as *variables* and *atomic data* have also been used for the data element concept.¹⁰⁸ The International Organization for Standardization (ISO)—a large SDO—characterizes data elements as the fundamental units of data that an organization manages and are necessarily part of the design of databases and of all data communicated to other organizations:

When an organization needs to transfer data to another organization, data elements are the fundamental units that make up the message. Messages occur between databases, between

¹⁰⁰ P.L. 107-347, §212(d); 116 Stat. 2941.

¹⁰¹ P.L. 107-347, §2(b)(3); 116 Stat. 2901.

¹⁰² P.L. 107-347, §101(a); 116 Stat. 2904; 44 U.S.C. §3602(8)(A).

¹⁰³ P.L. 104-13, §2; 109 Stat. 167; 44 U.S.C. §3504(b)(2).

¹⁰⁴ In 2005, OMB issued a congressionally mandated report on its implementation of a section of the E-Government Act that was concerned with the interoperability of federal systems and data integration. OMB described the FEA as addressing the goals of that particular section of the act (see OMB, *Report to Congress on Implementation of Section 212 of the E-Government Act of 2002*, December 17, 2005, p. 3, https://web.archive.org/web/20060124094500/http://www.whitehouse.gov/omb/egov/documents/Section_212_Report_Final.pdf#page=5). In 2013, OMB issued the “Federal Enterprise Architecture Framework v2” (see https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/fea_v2.pdf).

¹⁰⁵ The FEA consists of five interrelated reference models, one of which is the data reference model. For more, see OMB, *Federal Enterprise Architecture Framework v2*, January 29, 2013, p. 35, https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/fea_v2.pdf#page=35.

¹⁰⁶ Testimony of Randolph Hite in U.S. Congress, House Committee on Government Reform, Subcommittee on Technology, Information Policy, Intergovernmental Relations and the Census, *Federal Enterprise Architecture: A Blueprint for Improved Federal IT Investment Management and Cross-Agency Collaboration and Information Sharing*, 108th Cong., 2nd sess., May 19, 2004, committee print, Serial No. 108-227, p. 21, <https://www.govinfo.gov/content/pkg/CHRG-108hhrg96944/pdf/CHRG-108hhrg96944.pdf#page=25>.

¹⁰⁷ Examples of the term *data element* appearing in the *U.S. Code* include Title 5, Section 552a(o)(C), which concerns matching of records for the purposes of a “matching agreement” in a *matching program*; in Title 19, Section 1411(d)(2)(A), which specifies that an interagency steering committee must define, review, and update as necessary a standard set of data elements for the purposes of the National Customs Automation Program; and in Title 20, Section 1087nn(b), which concerns expected family contributions for the purposes of student financial assistance.

¹⁰⁸ As it relates to the term *variables*, see, for example, Executive Order 13985, Section 9: “Many Federal datasets are not disaggregated by race, ethnicity, gender, disability, income, veteran status, or other key demographic variables. This lack of data has cascading effects and impedes efforts to measure and advance equity” (available at <https://www.govinfo.gov/content/pkg/FR-2021-01-25/pdf/2021-01753.pdf>). As it relates to the term *atomic data*, see, for example, Center for Government Excellence at Johns Hopkins University, “OD [Open Data] Standards Definition,” <https://datastandards.directory/glossary#glossary-definition>.

databases and humans, and between humans. Moreover, the structure of databases don't have to be the same across organizations. So, the common unit for transferring data and related information is the data element.¹⁰⁹

Data element standards establish how to represent data, especially when there is an opportunity for confusion.¹¹⁰ These standards may, for example, establish definitions that should be consistently applied, constrain the data type to integer or date, require dollars and cents (\$4.60) versus whole dollars (\$5), or require only FIPS codes for geographic entities. As such, data element standards are comparable with the definition of *technical standards* in OMB's *Circular A-119* by establishing common and repeated rules for the production of data and by establishing definitions and other specifications.¹¹¹

Examples of Data Element Standards in Federal Policy and Their Implementation

- The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 required the Department of Health and Human Services (HHS) to define data elements for certain reports that states must submit on the Temporary Assistance for Needy Family (TANF) program.¹¹² The Administration for Children and Families (ACF), which administers TANF, noted that the information must be comparable and reliable in order to comply with the law's requirements, including reporting to Congress, and that "unless the reported data meet certain standards, [ACF] cannot adequately meet [its] responsibilities under the law."¹¹³ In effect, ACF promulgated standard data definitions that correspond to program-specific terms, such as *family*.¹¹⁴
- The Grant Reporting Efficiency and Agreements Act of 2019 requires data standards for federal grants and other financial awards, including definitions for data elements and standards that render reported information machine-readable.¹¹⁵ A GAO report indicates that definitions for 540 data elements related to government-wide awardmaking have been developed but are not fully consistent with the act's statutory requirements.¹¹⁶ Specifically, additional work is needed to specify the formatting requirements that allow the data to be machine-readable and consistently processed (e.g., California could be reported as "California," "Cali.," "CA," or "Ca"). GAO also found that certain data element definitions were not consistent with several "leading practices" for developing data definitions, including some that were ambiguous, which may increase uncertainty among federal award recipients when reporting data and result in inconsistent and noncomparable federal data.¹¹⁷

Semantic Standards

Sometimes, the term *semantic standards* is used to ensure that data elements are consistently and accurately interpreted by humans and computers.¹¹⁸ Some have characterized the term as being

¹⁰⁹ ISO, "Information Technology—Metadata Registries (MDR)—Part I: Framework," ISO/IEC 11179-1:2023, January 2023, p. 15. ISO standards are copyright protected.

¹¹⁰ Center for Government Excellence, "OD [Open Data] Standards Definition."

¹¹¹ OMB, *Circular A-119*, p. 15.

¹¹² P.L. 104-193, §411(a)(6); 110 Stat. 2150; 42 U.S.C. §611(7).

¹¹³ ACF, "Temporary Assistance for Needy Families Program (TANF)," 64 *Federal Register* 17858-17859, April 12, 1999, <https://www.govinfo.gov/content/pkg/FR-1999-04-12/pdf/99-8000.pdf#page=139>.

¹¹⁴ 45 C.F.R. §265.2.

¹¹⁵ P.L. 116-103, §4; 133 Stat. 3268; 31 U.S.C. §6402(a)(3)(A); 31 U.S.C §6402(c)(1).

¹¹⁶ See the "Highlights" section of GAO, *Grants Management: Action Needed to Ensure Consistency and Usefulness of New Data Standards*.

¹¹⁷ GAO, *Grants Management: Action Needed to Ensure Consistency and Usefulness of New Data Standards*, pp. 19-20, 24-26.

¹¹⁸ Panos Alexopoulos, *Semantic Modeling for Data: Avoiding Pitfalls and Breaking Dilemmas* (Sebastopol, CA: O'Reilly Media, 2020), p. 4.

weakly defined while also noting some overlap with the concept of metadata,¹¹⁹ which is discussed separately below. Additionally, organizations that develop what could be considered semantic standards do not necessarily identify their standards as such.¹²⁰ Nevertheless, in general, the function of semantic standards is to provide a language for data that is relevant for the domain in which the data are used.¹²¹ While some types of data standards, such as XML, can usually be implemented without a specific dependency on domain, semantic data standards, in contrast, represent not only “a way of looking at the world” but an agreement among groups about what the world looks like.¹²²

Semantic data standards have been characterized as enabling data interoperability and automating processes and transactions among data users.¹²³ Some observers note that interoperability standards have largely enabled technical aspects, with less attention paid to the semantic aspects.¹²⁴ Moreover, while semantic standards have been characterized as critical for enabling “true” interoperability,¹²⁵ it may be difficult for users to search for, identify, and reuse what is available.¹²⁶

While some laws refer to the semantic meaning of data, this is uncommon.¹²⁷ Some argue that unless semantic standards and specifications are identified, aligned, documented, managed, and promoted for reuse, the result is wasted investments in IT, which is more acute for governments given the size of their IT investments. This argument follows from the observation that efforts by other countries to use semantic standards for government operations and public administration have not produced widely accepted agreements on semantics for fundamental concepts.¹²⁸

Metadata and Metadata Standards

In practice, metadata is often simplified to mean “data about data”—for example, a column heading in a spreadsheet that describes the contents in the cells underneath the heading. The ISO explains that the usefulness of data for sharing is dependent on its meaning, type, format, and structure—all of which are metadata—being known to data users.¹²⁹

¹¹⁹ Vassilios Peristeras, “Semantic Standards: Preventing Waste in the Information Industry,” *IEEE Intelligent Systems*, vol. 28, no. 4 (July/August 2013), pp. 72-73.

¹²⁰ Peristeras, “Semantic Standards,” p. 73.

¹²¹ Boris Otto, Erwin Folmer, and Verena Ebner, “A Characteristics Framework for Semantic Information Systems Standards,” *Information Systems and e-Business Management*, vol. 10, no. 4 (December 2012), p. 576.

¹²² Peristeras, “Semantic Standards,” p. 72.

¹²³ M. Lynne Markus, Charles W. Steinfeld, and Rolf T. Wigand, “Industry-Wide Information Systems Standardization as Collective Action: The Case of the U.S. Residential Mortgage Industry,” *MIS Quarterly*, vol. 30 (August 2006), p. 453.

¹²⁴ Lampathaki et al., “Business to Business Interoperability,” p.1046. See also Peristeras, “Semantic Standards,” p. 72.

¹²⁵ Alexopoulos, *Semantic Modeling for Data*, p. 8. See also Erwin Folmer, Paul Oude Luttighuis, and Jos van Hillegersberg, “Do Semantic Standards Lack Quality: A Survey Among 34 Semantic Standards,” *Electronic Markets*, vol. 21 (June 2011), p. 100; and Lampathaki et al., “Business to Business Interoperability,” p. 1046.

¹²⁶ Peristeras, “Semantic Standards,” p. 73. A similar point is made in Lampathaki et al., “Business to Business Interoperability,” p. 1046, which characterizes the diversity of standards as a dilemma that makes achieving “true” interoperability more challenging.

¹²⁷ The idea of “semantic meaning” is included in the statutory definition of *machine-readable* that was discussed in the earlier section on “Data Format Standards” and is also used in the Financial Data Transparency Act of 2022 (P.L. 117-263, §5811(a); 136 Stat. 3422; 12 U.S.C. §5334(c)(1)(B)(ii)).

¹²⁸ Peristeras, “Semantic Standards,” pp. 72-73.

¹²⁹ ISO, “Information Technology—Metadata Registries (MDR)—Part I: Framework,” p. 28. ISO standards are copyright protected.

Metadata standards may establish specific rules for creating and managing metadata. In the same way there is variation in data standards, there are differences in metadata standards. Metadata standards may prescribe how to define data elements.¹³⁰ Some metadata standards are designed for publishing data on the web, including for producing machine-readable metadata.¹³¹ Metadata standards may provide “controlled vocabularies,” which control the actual values the metadata may take, and “content standards,” such as what specific metadata to record.¹³² Metadata standards may contribute to the quality of the underlying data.¹³³

Examples of Metadata in the Federal Context

- The OPEN Government Data Act defined *metadata* as “structural or descriptive information about data such as content, format, source, rights, accuracy, provenance, frequency, periodicity, granularity, publisher or responsible party, contact information, method of collection, and other descriptions.”¹³⁴ Some subsequent laws use the same definition.¹³⁵ The Geospatial Data Act of 2018 defined *metadata for geospatial data*.¹³⁶
- OMB advises agencies to “collect or create information in a way that supports downstream interoperability among information systems and streamlines dissemination to the public, where appropriate, by creating or collecting all new information electronically by default, in machine-readable open formats, using relevant data standards, that upon creation includes standard extensible metadata in accordance with OMB guidance.”¹³⁷
- A Senate committee recently heard testimony that AI can create links among datasets by using metadata in ways that may have been more difficult in the past, which may improve how the federal government delivers services.¹³⁸

Implementing Data Standards for Federal Data

Congress has sometimes directed specific agencies to play a government-wide role in data standards and guide and oversee their use in the executive branch. This section discusses the ways

¹³⁰ See GAO, *DATA Act: Data Standards Established, but More Complete and Timely Guidance Is Needed to Ensure Effective Implementation*, GAO-16-261, January 2016, pp. 11-12, <https://www.gao.gov/assets/gao-16-261.pdf#page=16>. In this report, GAO evaluated well-constructed data definitions, pointing to “leading practices” from ISO/IEC 11179-4—which is a standard concerned with creating and managing metadata—and evaluated the extent to which OMB and the Treasury Department developed definitions for financial reporting that were consistent with those leading practices.

¹³¹ Annette Griener et al., “Data on the Web Best Practices,” entry for Metadata, World Wide Web Consortium, January 31, 2017, <https://www.w3.org/TR/dwbp/#metadata>. In addition to discussing standardized metadata terms, these best practices also discuss standardized vocabularies (see entry for Data Vocabularies at the site listed above).

¹³² Jenn Riley, *Understanding Metadata: What Is Metadata, and What Is It For?* (Baltimore, MD: National Information Standards Organization, 2017), pp. 17-18, <https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf#page=20>.

¹³³ Bruce Bargmeyer and Daniel Gillman, “Metadata Standards and Metadata Registries: An Overview,” U.S. Bureau of Labor Statistics, 2000, p. 9, <https://www.bls.gov/osmr/research-papers/2000/pdf/st000010.pdf#page=9>.

¹³⁴ P.L. 115-435, §202(a); 132 Stat. 5535; 44 U.S.C. §3502(19).

¹³⁵ See, for example, P.L. 117-263, §5811(a); 136 Stat. 3422; 12 U.S.C. §5334(a)(2).

¹³⁶ P.L. 115-254, §752; 132 Stat. 3415; 43 U.S.C. §2801(11). *Metadata for geospatial data* means “information about geospatial data, including the content, source, vintage, accuracy, condition, projection, method of collection, and other characteristics or descriptions of the geospatial data.”

¹³⁷ OMB, “Revision of OMB Circular A-130, ‘Managing Information as a Strategic Resource,’” 81 *Federal Register* 49689, July 28, 2016, p. 15, https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf.

¹³⁸ Testimony of Beth Blauer, in U.S. Congress, Senate Committee on Homeland Security and Governmental Affairs, *Harnessing AI to Improve Government Services and Customer Services*, hearing, 118th Cong., 2nd sess., January 10, 2024.

OMB and NIST provide guidance to agencies on using data standards and the evolution of their roles.

Office of Management and Budget

OMB's role in providing guidance to agencies on data standards for federal data has evolved as the federal government has sought to capitalize on IT. For example, in September 1967, OMB (then known as the Bureau of the Budget) issued government-wide policies for standardizing data elements, in part to increase the federal government's ability to leverage automatic data processing—the term that was largely replaced by *information technology*.¹³⁹ OMB acknowledged at the time that computers and other technology were expanding opportunities for data integration, data aggregation, and data exchange but that the value of this data use could not be realized unless there were uniform understandings of data and the development and application of data standards.¹⁴⁰

As discussed below, OMB situates data standards in its guidance on information management, in statistical policy directives, and in a memorandum related to a federal data strategy. While OMB's guidance on IRM and federal statistical policy stems from its authority under the PRA and other statutes, the federal data strategy was initially a component of the Trump Administration's President's Management Agenda, which established a cross-agency priority goal to leverage data as a strategic asset.¹⁴¹

Data Standards for Information Resource Management

OMB issues guidance on “standards,” as a generic term, to implement IRM policies.¹⁴² A Senate committee report stated that IRM in the PRA was envisioned to include “IRM policy, IRM utilization to minimize paperwork burden, information dissemination, statistics, records management, information security and privacy, and IT management.”¹⁴³ Among other IRM functions, the administrator of the Office of Information and Regulatory Affairs—as delegated by the director of OMB—is tasked with:

¹³⁹ The legislative history indicates that the meaning of *automatic data processing* has evolved to mean IT for federal purposes (P.L. 104-106, §5602(b); 110 Stat. 699).

¹⁴⁰ Executive Office of the President, Bureau of the Budget, *Circular No. A-86. Standardization of Data Elements and Codes in Data Systems*, September 30, 1967, in Department of Commerce, National Bureau of Standards, *Federal Information Processing Standards Index*, January 1, 1971, p. 35, <https://files.eric.ed.gov/fulltext/ED048904.pdf#page=37>.

¹⁴¹ Executive Office of the President, *President's Management Agenda*, March 20, 2018, pp. 17-19, https://trumpadministration.archives.performance.gov/PMA/Presidents_Management_Agenda.pdf#page=17. The GPRA Modernization Act of 2010 requires OMB to develop “federal government priority goals” (P.L. 111-352, §5; 124 Stat. 3873; 31 U.S.C. §1120(a)(1)). In practice, these are called cross-agency priority goals. GAO states that each Administration typically releases a President's management agenda that communicates and organizes these goals and implementation strategies (see GAO, *Government Performance Management: Actions Needed to Improve Transparency of Cross-Agency Priority Goals*, GAO-23-106354, April 2023, p. 1, <https://www.gao.gov/assets/d23106354.pdf#page=5>). For background, see CRS Report R42379, *Changes to the Government Performance and Results Act (GPRA): Overview of the New Framework of Products and Processes*, by Clinton T. Brass.

¹⁴² P.L. 104-13, §2; 109 Stat. 167; 44 U.S.C. §3504(b)(1).

¹⁴³ U.S. Congress, Senate Committee on Governmental Affairs, *Paperwork Reduction Act of 1995*, report to accompany S. 244, 104th Cong., 1st sess., S.Rept. 104-8, February 14, 1995, p. 13, <https://www.congress.gov/104/crpt/srpt8/CRPT-104srpt8.pdf#page=16>.

1. developing and overseeing uniform IRM policies, principles, standards, and guidelines,¹⁴⁴ and
2. developing and using common standards for information collection, storage, processing and communication to foster greater federal data sharing, including standards for security, interconnectivity, and interoperability.¹⁴⁵

OMB issues guidance to agencies that specifically includes ways that data standards could be used for IRM.¹⁴⁶ For example, in *Circular A-130*, OMB advises agencies on using data standards to manage information, including:

- using open data standards to the maximum extent possible when implementing IT systems,
- standards that enable data governance, and
- using data standards to support downstream interoperability among information systems and to streamline dissemination of information to the public.¹⁴⁷

Data Standards for Federal Statistics

OMB oversees federal statistical data management through statistical policy directives.¹⁴⁸ Some of these directives establish specific definitions and other specifications for federal data and operate as data standards in this way. There is no single approach to how the data standards under these directives are developed. While largely for use by federal statistical agencies, in some cases these data standards are used for other nonstatistical purposes, including nonfederal ones, and to manage certain data, such as that collected from administrative forms.

For example, one of OMB's statistical directives establishes the North American Industry Classification System (NAICS) for classifying business establishments by their type of economic activity and is to ensure that business establishment data across the federal statistical system are

¹⁴⁴ 44 U.S.C. §3504(a)(1)(A). The PRA specifically requires the director to delegate to the administrator the authority to administer the functions that are contained within the act (44 U.S.C. §3503(b)).

¹⁴⁵ 44 U.S.C. §3504(b)(2). The PRA does not define *interoperability* for its purposes, although a committee report on the PRA makes several references to interoperability (U.S. Congress, Senate Committee on Governmental Affairs, *Paperwork Reduction Act of 1995*, report to accompany S. 244, 104th Cong., 1st sess., S.Rept. 104-8, February 14, 1995, <https://www.congress.gov/104/crpt/srpt8/CRPT-104srpt8.pdf>). For example, (1) “the bill maximizes utility by placing an emphasis on interoperability of agency systems and improvements in data sharing. These steps are meant to capitalize on the advantages that information technologies offer for streamlining agency operations, enhancing public access to government information, and reducing burdens on the public” (p. 23); (2) “an additional new purpose of the bill is to strengthen the partnership between the Federal government and State, local and tribal governments by minimizing information collection burdens and maximizing the utility of information collected by Federal agencies. This will require additional attention be paid to establishing common standards for data exchange and for interoperability among systems” (p. 24); and (3) “Developing interoperability among statistical systems in the different agencies also is important for improving access to valid and current data. To facilitate this coordination, the bill requires OMB to establish an interagency council, headed by the Chief Statistician and consisting of the heads of the major statistical agencies and representatives of other statistical agencies under rotating membership” (p. 27).

¹⁴⁶ For example, OMB issued guidance to agencies on metadata standards in conjunction with Executive Order 13642, describing it as helping “institutionalize the principles of effective information management at each stage of the information’s life cycle to promote interoperability and openness. Whether or not particular information can be made public, agencies can apply this framework to all information resources to promote efficiency and produce value” (OMB, *M-13-13*, p. 1).

¹⁴⁷ OMB, *Circular A-130*, pp. 8 and 15.

¹⁴⁸ 44 U.S.C. §3504(e)(3). See also CRS Insight IN12197, *The Federal Statistical System: A Primer*, by Taylor R. Knoedl.

comparable and can be aggregated for analysis.¹⁴⁹ NAICS is jointly developed by Mexico's Instituto Nacional de Estadística y Geografía (National Institute of Statistics and Geography), Statistics Canada, and OMB through the interagency Economic Classification Policy Committee. The Bureau of Labor Statistics uses NAICS for the Current Employment Statistics program, which is used to generate the monthly Employment Situation Summary (known more commonly as “the monthly jobs report”).¹⁵⁰ For nonstatistical purposes, the Small Business Administration uses NAICS to determine what is a “small business.”¹⁵¹

Another statistical policy directive establishes the Standard Occupational Classification (SOC) for categorizing all occupations in the U.S. national economy for which work is performed for pay or profit in the public, private, and military sectors.¹⁵² SOC is developed by a federal interagency technical working group. OMB states that SOC promotes a common language and encourages state and local governments to adopt SOC for classifying and analyzing occupations. At least one state requires SOC codes for employers' quarterly unemployment insurance reports.¹⁵³

Data Standards and a Federal Data Strategy

In 2019, OMB developed a federal data strategy and issued guidance to agencies on data management activities in federal programs and statistical programs and in support of agency missions.¹⁵⁴ The strategy was described as “a framework of operational principles and best practices that help agencies deliver on the promise of data in the 21st century.” The strategy has also been characterized as a vision for achieving a data-driven federal government by 2030.¹⁵⁵ Among other practices, the strategy called for adopting, adapting, creating as needed, and implementing data standards within relevant communities of interest in order to maximize data quality and facilitate data use, access, sharing, and interoperability.¹⁵⁶

OMB directed agencies to adhere to requirements within government-wide action plans it would develop.¹⁵⁷ Action plans were developed for 2020 under the Trump Administration¹⁵⁸ and 2021 under the Biden Administration,¹⁵⁹ with each establishing specific milestones for agencies to meet and timelines for doing so. OMB said it would assess agency progress through its existing oversight and coordination mechanisms, such as those stemming from the budget development

¹⁴⁹ OMB, “North American Industry Classification System—Revision for 2022; Update of Statistical Policy Directive No. 8, North American Industry Classification System: Classification of Establishments; and Elimination of Statistical Policy Directive No. 9, Standard Industrial Classification of Enterprises,” December 21, 2021, 86 *Federal Register* 72278, <https://www.govinfo.gov/content/pkg/FR-2021-12-21/pdf/2021-27536.pdf#page=2>. In practice, this may be referred to more simply as *Statistical Policy Directive No. 8*.

¹⁵⁰ See Department of Labor, Bureau of Labor Statistics, “Employment Situation Technical Note,” <https://www.bls.gov/news.release/empsit.tn.htm>.

¹⁵¹ 13 C.F.R. §121.101.

¹⁵² OMB, “Standard Occupational Classification (SOC) System—Revision for 2018,” November 28, 2017, 82 *Federal Register* 56272, <https://www.govinfo.gov/content/pkg/FR-2017-11-28/pdf/2017-25622.pdf#page=2>. In practice, this may be referred to more simply as *Statistical Policy Directive No. 10*.

¹⁵³ Louisiana uses SOC codes in this way. See La. R.S. 23:1660(A)(2).

¹⁵⁴ OMB, *M-19-18*.

¹⁵⁵ Executive Office of the President, *Federal Data Strategy: 2021 Action Plan*, October 2021, pp. 1-2, <https://strategy.data.gov/assets/docs/2021-Federal-Data-Strategy-Action-Plan.pdf#page=4>.

¹⁵⁶ OMB, *M-19-18*, p. 5.

¹⁵⁷ OMB, *M-19-18*, p. 7.

¹⁵⁸ Executive Office of the President, *Federal Data Strategy: 2020 Action Plan*, December 2019, <https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf>.

¹⁵⁹ Executive Office of the President, *Federal Data Strategy: 2021 Action Plan*.

process, information collections under the PRA, and system of records notices under the Privacy Act of 1974.¹⁶⁰ The 2021 action plan was released in October of that year. It noted that it was a transition year for executive branch agencies under a new Administration with its own priorities, and it indicated that agencies had until the end of the calendar year to work on the plan's "aspirational milestones." In 2023, a report from a science and technology policy group suggested that the federal data strategy needs additional effort by OMB to implement.¹⁶¹

National Institute of Standards and Technology

Both executive and congressional actions have shaped the role of NIST in data standards for federal data. By executive order, President Richard Nixon transferred to the National Bureau of Standards (NBS, NIST's precursor) functions that had previously been performed by OMB as federal agencies began to use computers more widely.¹⁶² The transfer resulted in NIST assuming a direct role in developing guidance to federal agencies on standardizing data elements, using data dictionaries, and managing data.¹⁶³ For example, NBS identified types of data standards, established policies for different types of data standards, and defined various agency responsibilities for implementing data standards.¹⁶⁴

NIST has a statutory role in supporting agencies in using voluntary consensus standards, which were discussed earlier. It also issues guidance to agencies that is broadly concerned with federal IT management. This guidance can, at times, have implications for agencies' use of data standards.

Information Technology Guidance

The Computer Security Act of 1987 charged NIST with "developing standards, guidelines, and associated methods and techniques for computer systems."¹⁶⁵ The act was largely concerned with ensuring standards for government-wide computer security and the privacy of sensitive information in federal computer systems.¹⁶⁶ In this way, the act provided a framework for managing risks that accompany the federal use of IT.

¹⁶⁰ OMB, *M-19-18*, p. 7.

¹⁶¹ Eric Egan, "Reviving and Reimagining the Federal Data Strategy for Mission Success," Information Technology and Innovation Foundation, June 5, 2023, <https://itif.org/publications/2023/06/05/reviving-and-reimagining-the-federal-data-strategy-for-mission-success/>.

¹⁶² Executive Order 11717, "Transferring Certain Functions from the Office of Management and Budget to the General Services Administration and the Department of Commerce," 38 *Federal Register* 12315, May 9, 1973. NBS was renamed as NIST by the Omnibus Trade and Competitiveness Act of 1988 (P.L. 100-418, §5101; 102 Stat. 1427).

¹⁶³ For examples of some of this guidance, see NBS, *Guideline for Choosing a Data Management Approach*, December 11, 1984, <https://nvlpubs.nist.gov/nistpubs/Legacy/FIPS/fipspub110.pdf>; and Judith J. Newton, *Guide on Data Entity Naming Conventions*, October 1987, <https://www.govinfo.gov/content/pkg/GOVPUB-C13-94ab71a32c5fe6f2c61a6c3ba14c307a/pdf/GOVPUB-C13-94ab71a32c5fe6f2c61a6c3ba14c307a.pdf>.

¹⁶⁴ Department of Commerce, "Standards of Data Elements and Representations," 38 *Federal Register* 33484, December 5, 1973, <https://www.govinfo.gov/content/pkg/FR-1973-12-05/pdf/FR-1973-12-05.pdf#page=38>. See also NBS, *Standardization of Data Elements and Representations*, December 5, 1973, <https://www.govinfo.gov/content/pkg/GOVPUB-C13-014d6f2898d47a6721672caa87ff91e4/pdf/GOVPUB-C13-014d6f2898d47a6721672caa87ff91e4.pdf#page=6>.

¹⁶⁵ P.L. 100-235, §3(2); 101 Stat. 1725.

¹⁶⁶ Section 2(b) of P.L. 100-235 enumerates the specific purposes of the act, including (1) "assign to the National Bureau of Standards responsibility for developing standards and guidelines for Federal computer systems, including responsibility for developing standards and guidelines needed to assure the cost-effective security and privacy of sensitive information in Federal computer systems;" (2) "to provide for promulgation of such standards and (continued...)"

Subsequent to the act's passage, the Department of Commerce—in which NIST operates— withdrew NIST's guidance to agencies that concerned certain data standards. A *Federal Register* notice of the withdrawal stated that such data standards were critical to federal information processing systems at the time and required special emphasis to support cost-effective information processing.¹⁶⁷ The department stated that the data standards were obsolete, as the prevailing policy emphasis was now on a broader view of IT management.

NIST's role in the federal use of standards took further shape when the NTTAA was enacted. As discussed earlier, the NTTAA established a preference for agencies to use voluntary consensus standards developed in the private sector. The NTTAA requires NIST to coordinate and promote agencies' use of private sector standards.¹⁶⁸ Thus, when a law specifies for data standards to incorporate voluntary consensus standards, NIST may be best positioned to support agencies in implementing this kind of requirement.

In accordance with certain federal IT laws, NIST develops guidance to agencies for the efficient operation of federal computer systems and for the security and privacy of such systems.¹⁶⁹ This guidance implicates federal data management and data standards in some cases, such as the categorizing of data collected or maintained by or on behalf of a federal agency for information security purposes.¹⁷⁰ NIST also publishes various reports, including the “special publications” series that provides guidelines, technical specifications, and recommendations to agencies.¹⁷¹ These publications may inform, for example, how an agency might use metadata to manage data¹⁷² or data governance and standards to enable “big data” interoperability.¹⁷³ Additionally, NIST produces reliable standardized technical and scientific data—formally defined as *standard reference data* in statute—which it may copyright and sell to scientists, engineers, and the public.¹⁷⁴

guidelines;” (3) “to require establishment of security plans by all operators of Federal computer systems that contain sensitive information;” and (4) “to require mandatory periodic training for all persons involved in management, use, or operation of Federal computer systems that contain sensitive information.”

¹⁶⁷ Department of Commerce, “Standardization of Data Elements and Representations,” *57 Federal Register* 30116, July 8, 1992, <https://www.govinfo.gov/content/pkg/FR-1992-07-08/pdf/FR-1992-07-08.pdf#page=26>.

¹⁶⁸ P.L. 100-235, §12(a)(3); 110 Stat. 782.

¹⁶⁹ For example, NIST will cite its responsibilities under P.L. 104-106, §5131 (110 Stat. 687); see also 41 U.S.C. §1441(a)(1); P.L. 107-347, §§302-303 (116 Stat. 2957); 15 U.S.C. §278g-3(a)(2-3); and 40 U.S.C. §11331(a-b). See also NIST, “Compliance FAQs: Federal Information Processing Standards (FIPS),” <https://www.nist.gov/standardsgov/compliance-faqs-federal-information-processing-standards-fips>.

¹⁷⁰ For these standards, see NIST, *Standards for Security Categorization of Federal Information and Information Systems*, February 2004, p. 1, <https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf#page=5>.

¹⁷¹ See NIST, “Publications,” <https://csrc.nist.gov/publications>.

¹⁷² NIST has described metadata as enhancing control over information flow and supporting the enforcement of allowable information flows. NIST explains that “information flow control” regulates where information can travel within and between systems, in contrast to who can access the information, and without regard to subsequent access to that information. NIST, *Security and Privacy Controls for Information Systems and Organizations*, September 2020, pp. 28-30, <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r5.pdf#page=55>.

¹⁷³ For example, NIST described key topics in the “NIST Big Data Interoperability Framework” (NBDIF). The NBDIF describes what is needed to leverage and draw insights from “big data,” a term used to characterize the large amounts of data that are available in a “networked, digitized, sensor-laden, and information-driven world.” One volume of the NBDIF discusses “big data management” (SP-1500-1r2). Another volume of the NBDIF discusses technical standards for big data and gaps in those standards (SP-1500-7r2). Access to each NBDIF volume can be found at <https://www.nist.gov/itl/big-data-nist/big-data-nist-documents/nbdif-version-30-final>.

¹⁷⁴ This function is authorized by the Standard Reference Data Act (P.L. 90-396), as amended (see P.L. 114-329, §108; 130 Stat. 2987; 15 U.S.C. §§290-294f). The definition of *standard reference data* is “(A) either (i) quantitative (continued...)

Recent Developments

The William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021 tasked NIST with developing best practices for using datasets to train AI systems.¹⁷⁵ These best practices are to address the use of metadata standards that describe the data's origins (i.e., provenance), the intent behind their creation, their permissible uses, their descriptive characteristics (including populations that are included and excluded), and any other properties determined by NIST.

In July 2021, NIST announced a project to focus on data management and data protection in certain use cases.¹⁷⁶ The project is intended to apply a “zero trust” security approach and is supposed to result in recommendations for defining data classifications and rules and standardizing how to communicate such classifications and rules at scale within the financial, government, manufacturing, technology, and health care sectors. The project intends to leverage existing NIST guidance and standards, including a metadata schema to describe attributes of subjects that might attempt to access otherwise protected data.¹⁷⁷

Considerations for Congress

Data standards may contribute to the usability of federal data for a particular purpose, including their readiness for use and their quality. Data standards have some clear applications in federal data management, particularly for agency operations and program management. The interoperability of federal data may also depend on data standards in various ways. Congress may also benefit from the use of data standards for federal data to assist in its decisionmaking.

Lawmakers may continue to identify a role for data standards in achieving certain policy outcomes in future legislation and for specific programs, regulatory matters, and government-wide operations. When Congress seeks to require data standards for federal data, it may consider addressing (1) such standards within specific policies, (2) such standards government-wide, and (3) the specification of data standards in the law and in cases of federal data interoperability.

information related to a measurable physical, or chemical, or biological property of a substance or system of substances of known composition and structure; (ii) measurable characteristics of a physical artifact or artifacts; (iii) engineering properties or performance characteristics of a system; or (iv) one or more digital data objects that serve—(I) to calibrate or characterize the performance of a detection or measurement system; or (II) to interpolate or extrapolate, or both, data described in subparagraph (A) through (C) [as in original]; and (B) that is critically evaluated as to its reliability under 15 U.S.C. §290b (15 U.S.C. §290a(1)).” See also NIST, “Standard Reference Data,” <https://www.nist.gov/srd>.

¹⁷⁵ P.L. 116-283, §5301; 134 Stat. 4538; 15 U.S.C. §278h-1(f)(1).

¹⁷⁶ Karen Scarfone and Murugiah Souppaya, *Data Classification Practices: Facilitating Data-Centric Security Management*, NIST, p. 3, <https://www.nccoe.nist.gov/sites/default/files/legacy-files/data-classification-project-description-final.pdf>. NIST describes data-centric security management as aiming “to enhance the protection of information (data) regardless of where the data resides or who it is shared with. Data-centric security management necessarily depends on organizations knowing what data they have, what its characteristics are, and what security and privacy requirements it needs to meet so the necessary protections can be achieved.” Data-centric security management is part of “zero trust,” which NIST describes as a “cybersecurity paradigm” and means “a collection of concepts and ideas designed to minimize uncertainty in enforcing accurate, least privilege per-request access decisions in information systems and services in the face of a network viewed as comprised” (see Scott Rose et al., *Zero Trust Architecture*, NIST, August 2020, p. 4, <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf#page=13>).

¹⁷⁷ Rose et al., *Zero Trust Architecture*, p. 8. Specifically, it cites Paul A. Grassi et al., *Attribute Metadata: A Proposed Schema for Evaluating Federated Attributes*, NIST, January 2018, p. 1, <https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8112.pdf#page=11>.

Policy Options for Data Governance and Data Management

Congress may consider ensuring that data standards within a given policy are accompanied by formal mechanisms for data management. For example, legislation could require data governance functions, such as developing the policies and processes for using data standards, establishing leadership and accountability roles, and ensuring oversight. Congress has identified a role for GAO in reporting on certain government-wide data standards. Likewise, lawmakers could consider specifying a role for inspectors general in auditing data standards management within agencies. When considering leadership and accountability for data standards, Congress may specifically consider CDOs.

Designating a Role for Chief Data Officers

Lawmakers may want to designate a role for an agency's CDO in implementing data standards, including in policies that concern only a specific program or activity. Including the CDO may contribute to how such standards are used and how they cohere with agency-wide data governance and IRM policies.¹⁷⁸ Including the CDO may also affect implementation of data standards as they are specified by Congress in a law. The Federal CDO Council identified agency CDOs as having a role in developing a common language for data-related terms within an agency, including standards-related terminology.¹⁷⁹ Involving the CDO may mitigate inconsistent interpretations of concepts and terms, particularly if such terms do not have single definitive interpretations in practice.

Although Congress intended CDOs to centralize responsibility for federal data management, there may be practical constraints on a CDO's ability to oversee data standards within an agency. CDOs are tasked with performing many functions.¹⁸⁰ In 2023, a majority of CDOs in medium and small agencies did not identify the CDO role as their primary responsibility, meaning that these CDOs also have other official responsibilities within their agencies,¹⁸¹ and a majority of CDOs in these agencies also do not have full-time equivalents to support their CDO functions.¹⁸² Thus, when requiring data standards, Congress may consider providing funding to support data standards management specifically—and perhaps data management more generally—within agencies.

Government-Wide Policy Options for Managing Data Standards

Federal data management practices will likely continue to develop, which may have a bearing on federal data standards. Part of this development may coincide with the implementation of laws such as FEBPA, while some may occur as agencies operationalize their data strategies. OMB acknowledges data governance in reducing barriers to using AI, suggesting that agencies may need to emphasize how they manage their data.¹⁸³ Similarly, NIST acknowledges the role of data

¹⁷⁸ Consistent with Title 44, Section 3506(b), of the *U.S. Code*, OMB's *Circular A-130* directs agencies to have agency-wide data governance policies that "clearly establish the roles, responsibilities, and processes by which agency personnel manage information as an asset and the relationships among technology, data, agency programs, strategies, legal and regulatory requirements, and business objectives" (p. 9).

¹⁷⁹ Federal CDO Council, *CDO Playbook: Advancing the Federal Data Strategy*, 2021, p. 8, https://resources.data.gov/assets/documents/CDO_Playbook_2021.pdf#page=8.

¹⁸⁰ 44 U.S.C. §3520(c).

¹⁸¹ Federal CDO Council, *CDO Council—Summer 2023 Survey*, December 2023, p. 17, https://www.cdo.gov/assets/documents/cdoc_final_10_26_2023.pdf#page=17. Error! Hyperlink reference not valid.

¹⁸² Federal CDO Council, *CDO Council—Summer 2023 Survey*, p. 66.

¹⁸³ OMB, *M-24-10*, p. 11.

standards in AI and in “big data” and interoperability, also suggesting that data management will play a role in these emerging areas.¹⁸⁴ As these activities move forward, Congress may consider reconciling the relevant policy frameworks to specifically account for federal data standards management and consider the role of NIST in coordinating the federal use of data standards.

Policy Coordination

Data standards operate within multiple policy frameworks, including those governing technical standards and information resources. These policy frameworks may need further coordination. Congress may consider whether to more specifically account for how agencies use data standards that are not voluntary consensus standards, including those that could be unique to the federal government.

Voluntary consensus standards may be impractical for federal data in certain ways, which the NTTAA anticipates.¹⁸⁵ Some federal data may be unique to federal operations and have no private sector equivalent that would be subjected to standards. For example, GAO reported that OMB and HHS did not incorporate voluntary consensus standards into data standards required by the Grant Reporting Efficiency and Agreements Transparency Act of 2019 because doing so was “neither reasonable nor practicable.”¹⁸⁶ This accords with the law: The act specifies use of voluntary consensus standards to the extent reasonable and practicable.¹⁸⁷ Some believe that in some situations the federal government may have to set data standards—including for its own internal data sharing and to reuse data it collects and manages—and should not shy away from this role.¹⁸⁸

OMB has a management and oversight role in IRM policies—which includes data standards for federal data—and also in technical standards through its guidance to agencies and the reporting it receives from agencies. While OMB’s *Circular A-119* permits agencies to use technical standards that are not voluntary consensus standards, such as government-unique standards, there is minimal guidance for using these types of standards. This includes when there is an absence of voluntary consensus standards or when an agency decides not to use them. For example, NIST has reported on the absence of data standards for AI systems, including those for data analytics, data exchange, data accessibility, and data quality.¹⁸⁹ At the same time, OMB’s guidance encourages agencies to adopt voluntary consensus standards for AI “as appropriate and consistent with OMB Circular No. A-119, if applicable.”¹⁹⁰

Currently, agencies are expected to report their use of standards that are not voluntary consensus standards, including their use of government-unique standards, to OMB under the NTTAA.¹⁹¹ Congress may consider whether government-unique standards apply to federal data and, as such,

¹⁸⁴ As it relates to AI, see NIST, *U.S. Leadership in AI*, p. 13. As it relates to “big data,” see NIST, *NIST Big Data Interoperability Framework: Volume 7, Standards Roadmap* (version 3), October 2019, pp. 35-37, <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-7r2.pdf#page=36>.

¹⁸⁵ P.L. 104-113, §12(d)(3); 110 Stat. 783.

¹⁸⁶ GAO, *Grants Management: Action Needed to Ensure Consistency and Usefulness of New Data Standards*, p. 22.

¹⁸⁷ P.L. 116-103, §4(a); 133 Stat. 3268; 31 U.S.C. §6402(c)(3).

¹⁸⁸ Michal S. Gal and Daniel L. Rubinfeld, “Data Standardization,” *New York University Law Review*, vol. 94, no. 4 (October 2019), p. 769.

¹⁸⁹ NIST, *U.S. Leadership in AI*, p. 11.

¹⁹⁰ OMB, *M-24-10*, p. 6.

¹⁹¹ P.L. 104-113, §12(d)(3); 110 Stat. 783. The NTTAA requires reporting to OMB, and OMB, in turn, directs agencies to submit annual reports to NIST, which then reports the use of government-unique standards in lieu of voluntary consensus standards to OMB (OMB, *Circular A-119*, pp. 33-34).

whether agencies should be required to comply with reporting requirements to OMB. Congress may also consider what purpose is served through such a reporting regime and how submitting annual reports contributes to how agencies substantively use data standards and to federal data management.

Designating a Role for NIST

Congress may also consider NIST's role in data standards. Currently, NIST coordinates technical standards for federal agencies. Designating a specific role for NIST may echo its previous role in establishing policies and issuing guidance concerned with data standards and the management of data within agencies. One policy option is to require NIST to coordinate data standards for federal data, including establishing requirements and guidelines, in a way similar to its responsibilities for setting federal information system standards,¹⁹² of which OMB oversees the development and implementation.¹⁹³ The role of NIST could thus be viewed as centralizing a framework for data standards. Evolving the government-wide role of NIST in data standards, however, is separate from the potential impacts that organizational culture, staffing, or resourcing have on agencies' ability to effectively use data standards and other best practices in data management, as discussed above.

Policy Options for Specifying Data Standards in the Law

For policymakers, it can be a challenge to define the data standards expected for a particular policy. Some laws, such as the Financial Data Transparency Act of 2022, specifically define *data standard* (i.e., “a standard that specifies rules by which data is described and recorded”),¹⁹⁴ and some may even specify various requirements for the data standard, including some technical requirements.¹⁹⁵ While such specification may clarify ambiguities, over-specification may have unforeseen consequences in practice.¹⁹⁶ Technical specifications may pose challenges where data interoperability is a policy objective.

Implications for Data Interoperability

Specifying technical details for data standards may affect the interpretation of their requirements, which may affect a policy's implementation and outcomes. These effects may be more notable where data are moving beyond typical agency boundaries (e.g., data interoperability; data integration). First, some stakeholders believe that Congress's legislative approach should be more principles-based and less prescriptive, focusing on processes and goals for how federal data are used.¹⁹⁷ Second, the terminology for data standards in cases of data integration and data interoperability is “not as clear as most people and organizations seem to assume.”¹⁹⁸ These issues

¹⁹² 15 U.S.C. §278g-3(a)(1).

¹⁹³ 44 U.S.C. §3504(h)(1)(B).

¹⁹⁴ P.L. 117-263, §5811; 136 Stat. 3422; 12 U.S.C. §5334(a)(3).

¹⁹⁵ P.L. 117-263, §5811; 136 Stat. 3422; 12 U.S.C. §5334(c)(1)(B).

¹⁹⁶ For example, HHS claims that it cannot develop an improper payment estimate for the TANF program because statute permits HHS to collect only the data elements specified in Section 411 of the Social Security Act (42 U.S.C. §611) (see GAO, *COVID-19: Current and Future Federal Preparedness Requires Fixes to Improve Health Data and Address Improper Payments*, GAO-22-105397, April 2022, p. 325, <https://www.gao.gov/assets/d22105397.pdf#page=337>).

¹⁹⁷ GSA, “The U.S. Data Federation Framework,” <https://federation.data.gov/us-data-federation-framework/#the-data-federation-playbook>.

¹⁹⁸ Lampathaki et al., “Business to Business Interoperability,” p. 1047.

may present a challenge to the implementation of data standards, because it is a process that depends on people and organizations. A policy that relies heavily on technical specifications to enable data interoperability may miss an important observation:

The hardest interoperability problems often have nothing to do with making technologies work together or enabling data to flow across systems. The hardest problems arise in linking together entire business processes and workflows across otherwise uncoordinated people and organizations.¹⁹⁹

For example, the Bipartisan Budget Act of 2018 required data exchange standards for maternal, infant, and early childhood home visiting programs to improve interoperability.²⁰⁰ These data standards would operate in an environment where data are collected by state government agencies and local providers.²⁰¹ HHS, which oversees the programs, found that the data exchange standards required by the act would be useful for states by supporting program operations, including for coordinating services provided to program participants, and for program evaluation. Nonetheless, HHS also found that certain activities are needed *before* the development of the standards, including establishing data governance to set the goals for data interoperability, ensuring that state-level information systems can support data interoperability, identifying a method for matching data across information systems within a state, and funding the adoption and use of federal data exchange standards.²⁰² Thus, use of such data exchange standards might be subordinate to other issues at play in data interoperability.

Certain data standards do not exist in some instances. For example, NIST identified several gaps in the standards needed for its Big Data Interoperability Framework, including for metadata specification, sharing and exchanging data, data quality and integrity, and general data management.²⁰³

Relying on standards alone will probably be insufficient to address some of the barriers to data interoperability. Congress may consider addressing some of these barriers and non-technical issues, such as the goals of data interoperability, the governance and advisory structures necessary for achieving such outcomes, and oversight of the efforts to produce those results.

For example, the Geospatial Data Act of 2018 established an interagency data committee to lead the development, management, and operations of national data infrastructure for geospatial data, which is to ensure the interoperability and sharing capabilities of federal information systems and data.²⁰⁴ One of the duties of the committee is to establish and maintain geospatial data standards,²⁰⁵ with the law also delineating several requirements for such standards²⁰⁶ and requiring agencies covered by the law to use these standards.²⁰⁷ Additionally, the law created an advisory committee—with membership to be selected from groups involved in the geospatial

¹⁹⁹ Palfrey and Gasser, *Interop*, p. 52.

²⁰⁰ P.L. 115-123, §50606(a); 132 Stat. 230; 42 U.S.C. §711(h)(5).

²⁰¹ Ivy Pool, Paul Wormeli, and Daniel Stein, *Developing Data Exchange Standards for MIECHV Home Visiting Programs*, September 2020, p. 22, https://www.acf.hhs.gov/sites/default/files/documents/opre/miechv_des_listening_session_summary_03sep20.pdf#page=25.

²⁰² Pool, Wormeli, and Stein, *Developing Data Exchange Standards*, pp. 21-27 and 36-37.

²⁰³ NIST, *NIST Big Data Interoperability Framework: Volume 7*, pp. 36-37, <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-7r2.pdf#page=37>.

²⁰⁴ P.L. 115-254, §753; 132 Stat. 3415; 43 U.S.C. §2802; P.L. 115-254, §755(b)(1)(E); 132 Stat. 3421; 43 U.S.C. §2804(b)(1)(E). The act is Subtitle F, Title VII, of the FAA Reauthorization Act of 2018.

²⁰⁵ P.L. 115-254, §753(c)(3); 132 Stat. 3416; 43 U.S.C. §2802.

²⁰⁶ P.L. 115-254, §757; 132 Stat. 3423; 43 U.S.C. §2806.

²⁰⁷ P.L. 115-254, §759(a)(6); 132 Stat. 3425; 43 U.S.C. §2808(a)(6).

community—to provide recommendations on implementing the national data infrastructure.²⁰⁸ Specific oversight mechanisms included in the act are inspector general audits that address compliance by covered agencies with the geospatial data standards and are to be submitted to Congress²⁰⁹ and evaluation by OMB of covered agencies’ budget justifications relative to their reported progress to meet their responsibilities under the law.²¹⁰

Author Information

Natalie R. Ortiz
Analyst in Government Organization and
Management

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS’s institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.

²⁰⁸ P.L. 115-254, §§754(b)(1), 754(e)(1); 132 Stat. 3418-3419; 43 U.S.C. §§2803(b)(1), 2803(e)(1).

²⁰⁹ P.L. 115-254, §759(c)(1); 132 Stat. 3426; 43 U.S.C. §2808(c)(1).

²¹⁰ P.L. 115-254, §759(b)(4); 132 Stat. 3426; 43 U.S.C. §2808(b)(4). For more on the Geospatial Data Act of 2018, see CRS Report R45348, *The Geospatial Data Act of 2018*, by Peter Folger.