



## **Curated Data Skills Catalog**

**September 2020**

## Table of Contents

<b>OVERVIEW AND PURPOSE .....</b>	<b>3</b>
<b>DEVELOPMENT PROCESS .....</b>	<b>3</b>
<b>FEDERAL DATA LIFECYCLE AND ASSOCIATED SKILLS .....</b>	<b>4</b>
FEDERAL DATA LIFECYCLE .....	4
ROLES AND SKILLS .....	8
<i>Data roles and skills</i> .....	9
<i>Other roles and skills</i> .....	12
<b>APPENDIX A .....</b>	<b>13</b>
DEVELOPMENT OF THE FEDERAL DATA LIFECYCLE .....	13
<b>APPENDIX B .....</b>	<b>14</b>
CROSSWALK OF FDS PRACTICES AND FEDERAL DATA LIFECYCLE ROLES .....	14

DRAFT

## Overview and Purpose

Improving staff data skills will allow agencies to better harness the power of data. To help Agencies address this need, Action 13 of the Federal Data Strategy (FDS) [2020 Action Plan](#) called for the development of a curated data skills catalog (Catalog). This catalog offers a common vocabulary of data roles and skills in the Federal data ecosystem. It highlights examples of governmental learning opportunities to help agencies develop staff data competencies within the Federal Data Lifecycle. The Catalog can also help agencies develop training programs by highlighting existing government resources and curricula.

Increasing agency staff data capabilities supports the implementation of both the Evidence Act and related guidance<sup>1-3</sup> as well as the Executive Order on Maintaining American Leadership in Artificial Intelligence,<sup>4</sup> both of which require an assessment of staff capacity for various data-related functions.

The Catalog consists of two parts:

- **Federal Data Lifecycle and Associate Skills (included for review)**, providing a common vocabulary of the data roles and skills in the Federal data ecosystem.
- **Federal Learning Opportunities (upcoming)**, detailing examples of existing Federal data training and courses.

## Development Process

This Catalog is intended to help agencies develop competencies for managing data as a strategic asset and making data-driven decisions. The 2020 Action Plan specifically tasked the General Services Administration (GSA) with leading Action 13. A focus group of data experts from across 9 Federal agencies was convened to determine the significant roles in the data ecosystem within the Federal government, list skills affiliated with these key functions, and create the draft Catalog.<sup>5</sup>

The Catalog was released in multiple stages to allow useful content to be available as soon as it was ready rather than in a single, delayed release. The contents of the Catalog have initially been available to the Chief Data Officers (CDOs) of each agency as well as other relevant Federal communities as it was being built. This allowed Federal stakeholder input to be incorporated early in its design. Demand for data skills identified in other parts of the FDS, including Action 4, have also informed the priorities for how the Catalog was built. It is now being released to obtain non-Federal stakeholder input. The Curated Data Skills Catalog will be posted on [resources.data.gov](https://resources.data.gov) on November 30, 2020 and transferred to the CDO Council after December 31, 2020.

---

<sup>1</sup> [Foundations for Evidence-Based Policymaking Act of 2018, Pub. L. No. 115-435, 132 Stat. 5529.](#)

<sup>2</sup> OMB M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, July 10, 2019, found at: [whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf](https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf).

<sup>3</sup> The Evidence Act also requires a capacity assessment, which in part determines “the extent to which evaluation and research capacity is present within the agency to include personnel” and “the extent to which the agency has the capacity to assist agency staff and program offices to develop the capacity to use evaluation research and analysis approaches and data in the day-to-day operations.”

<sup>4</sup> <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>

<sup>5</sup> The focus group convened to create the Federal Data Lifecycle is described in Appendix A.

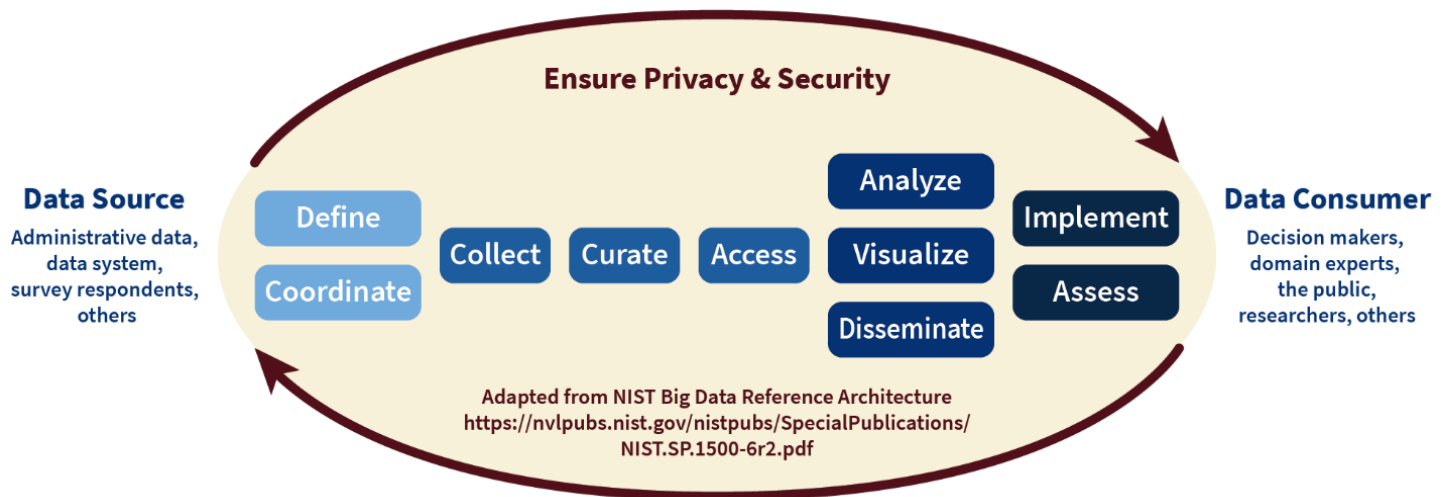
## Federal Data Lifecycle and Associated Skills

### Federal Data Lifecycle

This Lifecycle offers a common vocabulary for agencies to understand and evaluate data roles and skills. The Federal Data Lifecycle adapts the [NIST Big Data Reference Architecture](#)<sup>6</sup> and links subsequent data roles with the established [FDS practices](#). The framework below lays out a variety of roles for those involved in data and provides the conventional responsibilities for each role. These include data-focused roles<sup>7</sup> and others in the organization such as leadership and domain experts. This framework can be helpful with assessing staff data skills and creating training programs to bridge any gaps.

The required data skills, whether technical skills associated with analysis, ‘softer’ skills associated with communication and coalition building, or otherwise are given for each of the roles beginning on page 7. FDS practices are related to many of these data roles, and a crosswalk of FDS Practices and the Federal Data Lifecycle provides further details of this connection in Appendix B.

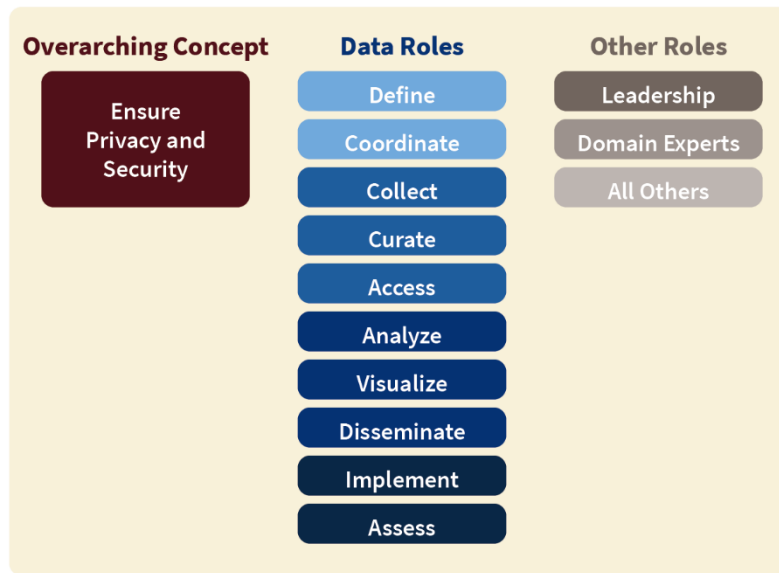
Figure 1: Federal Data Lifecycle



<sup>6</sup> U.S. Department of Commerce. National Institute of Standards and Technology (NIST). NIST Big Data Program. Retrieved from [nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-6r2.pdf](https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-6r2.pdf).

<sup>7</sup> Agency personnel in all data roles need to have a basic understanding of their agency's cybersecurity and privacy policies that govern data and need to work closely with their agency experts in cybersecurity and privacy throughout the Federal Data Lifecycle. More information on the Federal Information Security Modernization Act of 2014 can be found at [congress.gov/113/plaws/publ283/PLAW-113publ283.pdf](https://congress.gov/113/plaws/publ283/PLAW-113publ283.pdf)

Figure 2. Federal Data Lifecycle roles



#### Data roles

- **Define:** Identify agency and stakeholder needs for data of sufficient quality for intended uses
- **Coordinate:** Assess the ability of data resources and infrastructure to meet agency and stakeholder needs
- **Collect:** Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs
- **Curate:** Organize, refine, and maintain agency data resources with sufficient quality to meet agency and stakeholder needs
- **Access:** Identify and develop multiple data access methods for agency staff and stakeholders
- **Analyze:** Optimize the ability of staff and stakeholders to use agency data to generate insights
- **Visualize:** Present data insights for consumption by leaders and stakeholders
- **Disseminate:** Provide multiple avenues for release of data and insights
- **Implement & Assess:** Maximize the use of data for decision-making, accountability, and the public good and continuously improving the data process to address problems with data collection/measurement, processing, and analyses

#### Other roles

- **Leadership:** Cultivate and support a culture of enterprise-wide demand and use of data to maximize outcomes
- **All Others:** Each staff member values the use of data in their day-to-day work
- **Domain Expert:** Understands the context around the data, the needs of all involved stakeholders, and often takes on many roles throughout the process in an advisory or lead capacity to inform collection, data systems, data dictionaries, data set design, and analysis

- **Privacy and Security:** Ensure that agencies are consistently adopting and using the most up-to-date methods to protect data and comply with all applicable laws and regulations

DRAFT

**Select Statutory authority related to data**

- Privacy Act of 1974
- Federal Information Security Modernization Act of 2014 (FISMA) Public Law 113–283 (2014)
- Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act) – Public Law 115–435 (2019)
  - TITLE I—FEDERAL EVIDENCE–BUILDING ACTIVITIES
  - TITLE II—OPEN GOVERNMENT DATA ACT
  - TITLE III—CONFIDENTIAL INFORMATION PROTECTION AND STATISTICAL EFFICIENCY
- OMB Memorandum M–19– 15 Improving Implementation of the Information Quality Act issued on April 24, 2019. <https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf>
- Information Quality Act of 2001
- Geospatial Data Act of 2018

## Roles and Skills

Overarching Concept	Description
<p><b><u>Ensure Privacy and Security:</u></b> Ensure that agencies are consistently adopting and using the most up-to-date methods to protect data and comply with all applicable laws and regulations</p> <p><b><u>Practices</u></b></p> <ul style="list-style-type: none"> <li>11. Prioritize Data Governance</li> <li>12. Govern Data to Protect Confidentiality and Privacy</li> <li>13. Protect Data Integrity</li> </ul>	<p>Agency personnel in all data roles need to have a basic understanding of their agency's cybersecurity and privacy policies that govern data, and need to work closely with their agency experts in cybersecurity and privacy throughout the Federal Data Lifecycle</p>



Data roles and skills

Data Role	Skills
<p><b><u>Define:</u></b> Identify agency and stakeholder needs for data of sufficient quality for intended uses</p> <p><b><u>Practices</u></b></p> <ol style="list-style-type: none"> <li>1. Identify Data Needs to Answer Key Agency Questions</li> <li>2. Assess and Balance the Needs of Stakeholders</li> <li>28. Align Quality with Intended Use</li> </ol>	<ul style="list-style-type: none"> <li>• Strategic planning</li> <li>• Performance metrics</li> <li>• Strategic thinking</li> <li>• Problem solving</li> <li>• Planning</li> <li>• Communication</li> <li>• Teamwork</li> <li>• Knowledge of requirements for information collection review approvals</li> </ul> <p><i>Domain Expert input is important at this step</i></p>
<p><b><u>Coordinate:</u></b> Assess the ability of data resources and infrastructure to meet agency and stakeholder needs</p> <p><b><u>Practices</u></b></p> <ol style="list-style-type: none"> <li>1. Identify Data Needs to Answer Key Agency Questions</li> <li>8. Monitor and Address Public Perceptions</li> <li>9. Connect Data Functions Across Agencies</li> <li>10. Provide Resources Explicitly to Leverage Data Assets</li> <li>15. Assess Maturity</li> <li>16. Inventory Data Assets</li> <li>17. Recognize the Value of Data Assets</li> <li>18. Manage with a Long View</li> <li>19. Maintain Data Documentation</li> <li>20. Leverage Data Standards</li> <li>21. Align Agreements with Data Management Requirements</li> <li>36. Leverage Partnerships</li> <li>37. Leverage Buying Power</li> <li>38. Leverage Collaborative Computing Platforms</li> </ol>	<ul style="list-style-type: none"> <li>• Horizon scanning</li> <li>• Strategic thinking</li> <li>• Problem solving</li> <li>• Planning</li> <li>• Communication</li> <li>• Knowledge of the budget process</li> <li>• Business acumen</li> <li>• Knowledge of data standards</li> <li>• Federal enterprise architecture framework (FEAF) Data modelling</li> <li>• Relationship building</li> <li>• Coalition building</li> <li>• Knowledge of requirements for information collection review approvals</li> <li>• Knowledge of novel data sources and collection methods</li> </ul> <p><i>Domain Expert input is important at this step</i></p>
<p><b><u>Collect:</u></b> Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs</p> <p><b><u>Practices</u></b></p> <ol style="list-style-type: none"> <li>23. Allow Amendment</li> <li>25. Coordinate Federal Data Assets</li> <li>26. Share Data Between State, Local, and Tribal Governments and Federal Agencies</li> <li>29. Design Data for Use and Re-Use</li> <li>30. Communicate Planned and Potential Uses of Data</li> </ol>	<ul style="list-style-type: none"> <li>• Basic computer literacy</li> <li>• Collecting new source data such as through web scraping</li> <li>• Frame development, questionnaire/instrument development</li> <li>• Sampling design</li> <li>• Real-time or near real-time collection methods</li> <li>• Developing MOUs and Interagency Agreements</li> <li>• Communication</li> <li>• Interpersonal</li> <li>• Knowledge of requirements for information collection review approvals</li> <li>• Reviewing existing data for potential new uses</li> </ul>

Data Role	Skills
<p><b><u>Curate:</u></b> Organize, refine, and maintain agency data resources with sufficient quality to meet agency and stakeholder needs</p> <p><b><u>Practices</u></b>                      16. Inventory Data Assets                      19. Maintain Data Documentation                      20. Leverage Data Standards                      21. Align Agreements with Data Management Requirements                      24. Enhance Data Preservation                      32. Harness Safe Data Linkage                      33. Promote Wide Access</p>	<ul style="list-style-type: none"> <li>• Techniques to understand and clean raw and unstructured data (data wrangling)</li> <li>• Knowledge of operating systems</li> <li>• Develop scalable extract, transform, and load (ETL) processes</li> <li>• Big data technologies</li> <li>• Data warehouse/architecture design, development, construction, and maintenance</li> <li>• Programming tools</li> <li>• Statistics, data checking, and internal controls on data reporting and quality, consistency, and logic edits</li> <li>• Information quality control and quality assurance methods and techniques</li> <li>• Metadata skills including knowledge of the enterprise metadata standards and ability to produce unambiguous, useful metadata</li> <li>• Understand and incorporate stakeholder needs and requirements</li> </ul> <p><i>Domain Expert input is important at this step</i></p>
<p><b><u>Access:</u></b> Identify and develop multiple data access methods for agency staff and stakeholders</p> <p><b><u>Practices</u></b>                      5. Prepare to Share                      22. Identify Opportunities to Overcome Resource Obstacles                      25. Coordinate Federal Data Assets                      26. Share Data Between State, Local, and Tribal Governments and Federal Agencies                      31. Explicitly Communicate Allowable Use                      33. Promote Wide Access                      34. Diversify Data Access Methods                      35. Review Data Releases for Disclosure Risk</p>	<ul style="list-style-type: none"> <li>• Disclosure risk limitation</li> <li>• File formats</li> <li>• Business acumen</li> <li>• Communication</li> <li>• Data sharing policies</li> <li>• Database authentication methods</li> <li>• Tiered access to data methods</li> <li>• Understand stakeholders needs and requirements</li> <li>• Data modeling and transformation</li> <li>• Metadata repository development and maintenance</li> </ul>
<p><b><u>Analyze:</u></b> Optimize the ability of staff and stakeholders to use agency data to generate insights</p> <p><b><u>Practices</u></b>                      27. Increase Capacity for Data Management and Analysis                      31. Explicitly Communicate Allowable Use                      38. Leverage Collaborative Computing Platforms                      39. Support Federal Stakeholders                      40. Support Non-Federal Stakeholders</p>	<ul style="list-style-type: none"> <li>• Math and statistics</li> <li>• Trend or pattern analysis</li> <li>• Predictive modeling and projections</li> <li>• Geospatial analysis</li> <li>• Machine learning algorithms, artificial intelligence, natural language processing and rapid process automation</li> <li>• Data conditioning and data mining</li> <li>• Communication</li> <li>• Generate dashboards, charts, and graphs</li> <li>• Understand metadata</li> </ul> <p><i>Domain Expert input is important at this step</i></p>

Data Role	Skills
<p><b><u>Visualize:</u></b> Present data insights for consumption by leaders and stakeholders</p> <p><b><u>Practices</u></b> 6. Convey insights from data</p>	<ul style="list-style-type: none"> <li>• Programming or tools to create static or interactive reports, graphics, and dashboards</li> <li>• Web design and HTML</li> <li>• Data visualization (including 508 compliance)</li> <li>• Understand stakeholder needs and requirements</li> <li>• User experience design</li> <li>• Communication</li> <li>• Data storytelling</li> </ul>
<p><b><u>Disseminate:</u></b> Provide multiple avenues for release of data and insights</p> <p><b><u>Practices</u></b> 6. Convey Insights from Data 7. Use Data to Increase Accountability 14. Convey Data Authenticity 30. Communicate Planned and Potential Uses of Data 31. Explicitly Communicate Allowable Use 33. Promote Wide Access 34. Diversify Data Access Methods 35. Review Data Releases for Disclosure Risk</p>	<ul style="list-style-type: none"> <li>• Communication</li> <li>• Negotiation</li> <li>• Understand stakeholder needs and requirements</li> <li>• Relationship building</li> <li>• Data formats and API technology</li> <li>• Companion materials creation</li> <li>• Accessibility standards (508 compliance)</li> </ul>
<p><b><u>Implement &amp; Assess:</u></b> Maximize the use of data for decision-making, accountability, and the public good, and continuously improving the data process</p> <p><b><u>Practices</u></b> 2. Assess and Balance the Needs of Stakeholders 3. Champion Data Use 4. Use Data to Guide Decision-Making 7. Use data to Increase Accountability</p>	<ul style="list-style-type: none"> <li>• Strategic and innovative thinking</li> <li>• Problem solving</li> <li>• Data management</li> <li>• Scientific standards, theories, measurement, testing, and evaluation procedures</li> <li>• Communication</li> <li>• Coalition building and advocate for data use</li> </ul>

Other roles and skills

Other Roles	Skills
<p><b><u>Leadership</u></b> Cultivate and support a culture of enterprise-wide harmonization of data-related activities to maximize outcomes</p>	<ul style="list-style-type: none"> <li>• Data use</li> <li>• Quantitative and qualitative data types</li> <li>• Communication</li> <li>• Relationship building</li> <li>• Data science process</li> <li>• Data sets, data priorities, and resources</li> <li>• Data-based decision making</li> </ul>
<p><b><u>All Others</u></b> Each staff member values the use of data in their day-to-day work</p>	<ul style="list-style-type: none"> <li>• Data use and applications</li> <li>• Basic terminology</li> <li>• Data risks or limitations</li> <li>• Reading charts and graphs</li> </ul>
<p><b><u>Domain Expert</u></b></p> <ul style="list-style-type: none"> <li>• Understands the context around the data, the needs of all involved stakeholders, and often takes on many roles throughout the process in an advisory or lead capacity to inform collection, data systems, data dictionaries, data set design, and analysis</li> <li>• Prioritizes work to improve operational efficiency, mitigate bias and maximize effective outcomes</li> <li>• Provides a variety of insights to maximize data work (use), including knowledge of customers/beneficiaries, full understanding of processes and requirements, connections with a larger domain ecosystem of experts and peers, and more</li> <li>• Fully understands the landscape of the data ecosystem and the needs of all involved stakeholders</li> <li>• Acquires a complete understanding of all data assets and data-related business processes and their respective responsiveness to internal and external stakeholder needs/requirements</li> </ul>	<ul style="list-style-type: none"> <li>• Mission-related subject matter expertise</li> <li>• Operations and procedures</li> <li>• Data Lifecycle</li> <li>• Strategic thinking</li> <li>• Data collection and use</li> <li>• Understand stakeholder needs and requirements</li> <li>• Relationship building</li> <li>• Communication</li> <li>• Knowledge of the possible data uses and limitations</li> </ul>

## Appendix A

### Development of the Federal Data Lifecycle

The GSA convened a focus group consisting of leaders from the Federal CDO, Statistical Officer and Chief Human Capital Officer communities to determine the significant roles in the data ecosystem within the Federal government and list skills and responsibilities associated with these key functions. The Federal Data Lifecycle was generated by adapting the [NIST Big Data Reference Architecture](#) which is a vendor-neutral, technology- and infrastructure-agnostic conceptual model. The Federal Data Lifecycle and corresponding skills framework form the backbone for the Catalog.

Focus Members	Group	Role & Organization
Anne Levine		Deputy CDO, Federal Communications Commission
Brian Moyer		Director of National Center for Health Statistics & HHS Statistical Official
Chris Haffer		CDO, Equal Employment Opportunity Commission
Keith Krut		HR Division, National Aeronautics and Space Administration
Lance Harris		Statistical Official, United States Office of Personnel Management
Lucas Hitt		U.S. Bureau of Economic Analysis
Megan Dreher		Performance Team, Office of Management and Budget
Megan McNeely		HR Division, U.S. Census Bureau
Sharon Boivin		Chief of Staff, Office of the CDO, U.S. Department of Education

## Appendix B Crosswalk of FDS Practices and Federal Data Lifecycle Roles

	Overarching Concept	Data Lifecycle								
	Ensure Privacy & Security	Define	Coordinate	Collect	Curate	Access	Analyze	Visualize	Disseminate	Implement & Assess
	Ensure that agencies are consistently adopting and using the most up-to-date methods to protect data and comply with all applicable laws and regulations	Identify agency and stakeholder needs for data of sufficient quality for intended uses	Assess the ability of data resources and infrastructure to meet agency and stakeholder needs	Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs	Organize, refine, and maintain agency data resources with sufficient quality to meet agency and stakeholder needs	Identify and develop multiple data access methods for agency staff and stakeholders	Optimize the ability of staff and stakeholders to use agency data to generate insights	Present data insights for consumption by leaders and stakeholders	Provide multiple avenues for release of data and insights	Maximize the use of data for decision-making, accountability, and the public good and continuously improving the data process
1. Identify Data Needs to Answer Key Agency Questions: Use the learning agenda process to identify and prioritize the agency's key questions and the data needed to answer them.		✓	✓							
2. Assess and Balance the Needs of Stakeholders: Identify and engage stakeholders throughout the data lifecycle to identify stakeholder needs and to incorporate stakeholder feedback into government priorities to maximize entrepreneurship, innovation, scientific discovery, economic growth, and the public good.		✓								✓
3. Champion Data Use: Leaders set an example, incorporating data in decision-making and targeting resources to maximize the value of data for decision-making, accountability, and the public good.										✓
4. Use Data to Guide Decision-Making: Effectively, routinely, transparently, and appropriately use data in policy, planning, and operations to guide decision-making; share the data and analyses behind those decisions.										✓
5. Prepare to Share: Assess and proactively address the procedural, regulatory, legal, and cultural barriers to sharing data within and across federal agencies, as well as with external partners.						✓				
6. Convey Insights from Data: Use a range of communication tools and techniques to effectively present insights from data to a broad set of audiences.								✓	✓	
7. Use Data to Increase Accountability: Align operational and regulatory data inputs with performance measures and other outputs to help the public to understand the results of federal investments and to support informed decision-making and rule-making.									✓	✓
8. Monitor and Address Public Perceptions: Regularly assess and address public confidence in the value, accuracy, objectivity, and privacy protection of federal data to make strategic improvements, advance agency missions, and improve public messages about planned and potential uses of federal data.				✓						
9. Connect Data Functions Across Agencies: Establish communities of practice for common agency data functions (e.g. data management, access, analytics, informatics, and user support) to promote efficiency, collaboration, and coordination.				✓						
10. Provide Resources Explicitly to Leverage Data Assets: Ensure that sufficient human and financial resources are available to support data driven agency decision-making, accountability and the ability to spur commercialization, innovation, and public use.				✓						
11. Prioritize Data Governance: Ensure there are sufficient authorities, roles, organizational structures, policies, and resources in place to transparently support the management, maintenance, and use of strategic data assets.	✓									
12. Govern Data to Protect Confidentiality and Privacy: Ensure there are sufficient authorities, roles, organizational structures, policies, and resources in place to provide appropriate access to confidential data and to maintain public trust and safeguard privacy.	✓									
13. Protect Data Integrity: Emphasize state-of-the-art data security as part of information Technology security practices for every system that is refreshed, architected, or replaced to address current and emerging threats; foster innovation and leverage new technologies to maintain protection.	✓									
14. Convey Data Authenticity: Disseminate data sets such that their authenticity is discoverable and verifiable by users throughout the information lifecycle, consistent with open data practices, and encourage appropriate attribution from users.									✓	
15. Assess Maturity: Evaluate the maturity of all aspects of agency data capabilities to inform priorities for strategic resource investment.				✓						
16. Inventory Data Assets: Maintain an inventory of data assets with sufficient completeness, quality, and metadata to facilitate discovery and collaboration in support of answering key agency questions and meeting stakeholder needs.				✓		✓				
17. Recognize the Value of Data Assets: Assign value to data assets based on maturity, key agency questions, stakeholder feedback, and applicable law and regulation to appropriately prioritize and document resource decisions.				✓						

	Overarching Concept	Data Lifecycle								
	Ensure Privacy & Security	Define	Coordinate	Collect	Curate	Access	Analyze	Visualize	Disseminate	Implement & Assess
	Ensure that agencies are consistently adopting and using the most up-to-date methods to protect data and comply with all applicable laws and regulations	Identify agency and stakeholder needs for data of sufficient quality for intended uses	Assess the ability of data resources and infrastructure to meet agency and stakeholder needs	Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs	Organize, refine, and maintain agency data resources with sufficient quality to meet agency and stakeholder needs	Identify and develop multiple data access methods for agency staff and stakeholders	Optimize the ability of staff and stakeholders to use agency data to generate insights	Present data insights for consumption by leaders and stakeholders	Provide multiple avenues for release of data and insights	Maximize the use of data for decision-making, accountability, and the public good and continuously improving the data process
18. <b>Manage with a Long View:</b> Include data investments in annual capital planning processes and associated guidance to ensure appropriated funds are being used efficiently to leverage data as a strategic long-term asset.			✓							
19. <b>Maintain Data Documentation:</b> Store up-to-date and comprehensive data documentation in accessible repositories to facilitate use and document quality, utility, and provenance in support of informing key agency questions and meeting stakeholder needs.			✓		✓					
20. <b>Leverage Data Standards:</b> Adopt or adapt, create as needed, and implement data standards within relevant communities of interest to maximize data quality and facilitate use, access, sharing, and interoperability.			✓		✓					
21. <b>Align Agreements with Data Management Requirements:</b> Establish terms and conditions for contracts, grants, cooperative agreements, and other agreements that meet data management requirements for processing, storage, access, transmission, and disposition.			✓		✓					
22. <b>Identify Opportunities to Overcome Resource Obstacles:</b> Coordinate with stakeholders to identify mutually-acceptable cost recovery, shared service, or partnership opportunities to enable data access while conserving available resources to meet user needs.						✓				
23. <b>Allow Amendment:</b> Establish clear procedures to allow members of the public to access and amend federal data about themselves, as appropriate and in accordance with federal laws, regulations and policies, in order to safeguard privacy, reduce potential harm from inaccurate data, and promote transparency.				✓						
24. <b>Enhance Data Preservation:</b> Preserve federal data in accordance with applicable law, regulation, policy, approved schedules, and mission relevance.					✓					
25. <b>Coordinate Federal Data Assets:</b> Coordinate and share data assets across federal agencies to advance progress on shared and similar objectives, fulfill broader federal information needs, and reduce collection burden.				✓		✓				
26. <b>Share Data Between State, Local, and Tribal Governments and Federal Agencies:</b> Facilitate data sharing between state, local, and tribal governments and the Federal Government, where relevant and appropriate and with proper protections, particularly for programs that are federally funded and locally administered, to enable richer analyses for more informed decision-making.				✓		✓				
27. <b>Increase Capacity for Data Management and Analysis:</b> Educate and empower the federal workforce by investing in training, tools, communities, and other opportunities to expand capacity for critical data-related activities such as analysis and evaluation, data management, and privacy protection.							✓			
28. <b>Align Quality with Intended Use:</b> Data likely to inform important public policy or private sector decisions must be of appropriate utility, integrity, and objectivity.		✓								

	Overarching Concept	Data Lifecycle								
	Ensure Privacy & Security	Define	Coordinate	Collect	Curate	Access	Analyze	Visualize	Disseminate	Implement & Assess
	Ensure that agencies are consistently adopting and using the most up-to-date methods to protect data and comply with all applicable laws and regulations	Identify agency and stakeholder needs for data of sufficient quality for intended uses	Assess the ability of data resources and infrastructure to meet agency and stakeholder needs	Organize, plan, and execute data collections and acquisitions to meet agency and stakeholder needs	Organize, refine, and maintain agency data resources with sufficient quality to meet agency and stakeholder needs	Identify and develop multiple data access methods for agency staff and stakeholders	Optimize the ability of staff and stakeholders to use agency data to generate insights	Present data insights for consumption by leaders and stakeholders	Provide multiple avenues for release of data and insights	Maximize the use of data for decision-making, accountability, and the public good and continuously improving the data process
29. Design Data for Use and Re-Use: Design new data collections with the end uses and users in mind to ensure that data are necessary and of high enough quality to meet planned and future agency and stakeholder needs.				✓						
30. Communicate Planned and Potential Uses of Data: Review data collection procedures to update and improve how planned and future uses of data are communicated, promoting public trust through transparency.				✓					✓	
31. Explicitly Communicate Allowable Use: Regularly employ descriptive metadata that provides clarity about access and use restrictions for federal data, explicitly recognizes and safeguards applicable intellectual property rights, conveys attribution as needed, and optimizes potential value to stakeholders to maximize appropriate legal use.						✓	✓		✓	
32. Harness Safe Data Linkage: Test, review, and deploy data linkage and analysis tools that use secure and privacy-protective technologies to address key agency questions and meet stakeholder needs while protecting privacy.					✓					
33. Promote Wide Access: Promote equitable and appropriate access to data in open, machine-readable form and through multiple mechanisms, including through both federal and non-federal providers, to meet stakeholder needs while protecting privacy, confidentiality, and proprietary interests.					✓	✓			✓	
34. Diversify Data Access Methods: Invest in the creation and usability of multiple tiers of access to make data as accessible as possible while minimizing privacy risk and protecting confidentiality.						✓			✓	
35. Review Data Releases for Disclosure Risk: Review federal data releases to the public to assess and minimize the risk of re-identification, consistent with applicable laws and policies, and publish reviews to promote transparency and public trust.						✓			✓	
36. Leverage Partnerships: Create and sustain partnerships that facilitate innovation with commercial, academic, and other partners to advance agency mission and maximize economic opportunities, intellectual value, and the public good.			✓							
37. Leverage Buying Power: Monitor needs and systematically leverage buying power for private sector data assets, services, and infrastructure to promote efficiency and reduce federal costs.			✓							
38. Leverage Collaborative Computing Platforms: Periodically review and optimize the use of modern collaborative computing platforms to minimize costs, improve performance, and increase use.			✓				✓			
39. Support Federal Stakeholders: Engage with relevant agencies to share expert knowledge of data assets, promote wider use, improve usability and quality, and meet mission goals.							✓			
40. Support Non-Federal Stakeholders: Engage with industry, academic, and other non-federal users of data to share expert knowledge of data assets, promote wider use, improve usability and quality, and advance innovation and commercialization.							✓			