

The background of the slide is a photograph of the United States Capitol dome in Washington, D.C. The dome is white with a prominent central dome and a smaller dome on top. The sky is blue with some white clouds. The image is partially obscured by a large blue geometric shape on the left and a grey shape at the bottom.

CDO Council

Summary of RFI Responses

This report provides a high-level summary of feedback provided to the CDO Council's 2021 Request for Information.



Background	3
Section 1 - General Comments	4
Section 2 - Data Skills and Workforce	6
Section 3 - Data Inventory	7
Section 4 - Data Sharing	8
Section 5 - Value and Maturity	9
Section 6 - Ethics and Equity	10
Section 7 - Technology	11

Background

The Chief Data Officers Council

The Federal Chief Data Officers (CDO) Council was established by the Foundations for Evidence-Based Policymaking Act, which also requires all federal agencies to appoint a Chief Data Officer. The Council's vision is to improve government mission achievement and increase the benefits to the Nation through improvement in the management, use, protection, dissemination, and generation of data in government decision-making and operations.

The CDO Council is composed of over 80 CDOs from across the Federal Government, as well as representatives from the U.S. Office of Management and Budget, and other key councils and committees. Additionally, the Council has working groups and committees that focus on specific topics to help agencies connect and collaborate.

Public Meeting and Request for Information

On October 14, 2021, the CDO Council held a virtual public meeting. The meeting highlighted the Council's first year accomplishments, priorities, and working group updates. During the meeting, large and small agency CDOs shared information about their data activities, the Evidence Act Councils shared information regarding federal data collaboration efforts, and the CDO Council shared its [Request for Information](#) (RFI) for public input.

The RFI was published to obtain input, information, and recommendations from a broad array of public stakeholders on available methods, approaches, and tools to assist in the CDO Council's efforts and guide its focus areas.

RFI Review Process Overview

After [submissions](#) were received and processed, the CDO Council shared the responses to the RFI with the relevant working groups and other stakeholders in the Council. The reviewers for each section read the responses and identified key feedback and recurring themes used in developing this summary.

Section 1 - General Comments

Common Themes

Respondents to Section 1 of the RFI submitted topic suggestions for the CDO Council's focus areas and provided updates on industry trends in the data space. Common themes included:

- Coordinating with state, local, and tribal governments.
- Implementing a Federal Government-wide approach for data practices to increase data connectivity and transparency across agencies.

Several best practices and methodologies were provided for the CDO Council's awareness.

Key Feedback

Respondents stated that an opportunity is being missed to address comprehensive Master Data Management (MDM) at and across federal civilian agencies. They recommended that the CDO Council should add MDM as a focus area because exploring data management through this lens could address many critical challenges for government agencies and offer guidance on the capabilities needed in modern data management solutions. Respondents suggested that the right MDM approach can act as the glue that binds agency systems and information together creating a single source of truth and providing trusted, authoritative, and complete data that is consistent across programs, services, and agencies.

Respondents highlighted the following industry trends in data quality, analytics, and data frameworks for the CDO Council's awareness.

- Data quality: crowdsourced collaboration and real-time data integration.
- Outcomes with analytics: increased analytic rigor, operationalized analytics, and more granular security.
- Better frameworks: data sharing, data mesh, and multi-cloud data fabric.

Respondents emphasized that industry trends in data management and governance focus on a bottom-up approach by using data discovery to drive data insights in order to reduce manual effort and eliminate repetitive tasks. Respondents suggested that the CDO Council member agencies can leverage machine learning to automate the identification and governance of sensitive and critical information.

Respondents noted that the CDO Council and CDOs could play a critical role in supporting state, local, and tribal capacity for data collection, access, and use. Respondents suggested the Council should engage with CDO networks, potentially through a workgroup, to identify needs and best practices and understand barriers to data use at the state, local, and tribal levels and how federal agencies can support improved data capacity.

Respondents would like Federal Agencies to look to state, local, and tribal agencies as partners, rather than ordinary stakeholders and should be required to incorporate state, local, and tribal data and expertise into their collection, analysis, and decision-making processes. They stated that agencies should also consult with state, local, and tribal partners to establish uniform data standards and ensure access to timely and reliable federal datasets.

Respondents stated that the CDO Council should address data silos by identifying and defining contributors to siloed approaches to data use, making recommendations for innovative approaches to

collaboration across federal agencies to align requirements and definitions, and advocating within agencies for policies and practices that will result in improved collaboration.

Section 2 - Data Skills and Workforce

Common Themes

Respondents to Section 2 of the RFI provided several suggestions for improving data skills in the federal workforce. Common themes included:

- Improving data literacy across all job series (beyond purely technical roles).
- providing regular training specifically for enhancing data skills.
- improving hiring processes and retention rates.

Key Feedback

When thinking about data literacy, respondents suggested the Federal Government focus on overall improvement at the enterprise level, including the general workforce, day-to-day roles, and leadership roles.

Respondents recommended that Agencies:

- Upskill their workforce by identifying the level of data acumen needed by different groups in the organization (e.g., executives, managers, technical and non-technical staff).
- Define the roles and responsibilities of effective data science teams.
- Describe different pedagogical models available to upskill federal staff, including their pros and cons.
- Describe organizational culture features necessary to support staff data upskilling, including the upskilling of the non-technical staff and executive leadership.
- Describe organizational strategies to best attract and retain data scientists in government.
- Develop benchmarking standards for similar agencies to measure oneself against.

Respondents stated that agencies should set aside funds specifically for data skills improvement. They also noted that training on leading edge data technology, robotic process automation, artificial intelligence, and machine learning should be emphasized. Finally, respondents highlighted that normalized training that uses an organization's data to answer real challenges may also be beneficial as the training is tailored to the organization.

To improve hiring processes, respondents suggested agencies use direct-hire authorities, provide scholarships or other incentives, partner with academia/universities in curriculum building and to market federal roles in data areas, and enhance relationships with data-focused associations. Additionally, respondents emphasized that agencies should look beyond degrees alone and consider hiring individuals with desirable experience regardless of their educational background (including hiring staff from different disciplines as long as they have the requisite data skills).

Respondents noted agencies should consider using new hiring strategies, such as the Subject Matter Expert - Qualifications Assessment (SME-QA) to recruit data scientists for multiple agencies. Specifically, using flexible hiring authorities, such as the Intergovernmental Personnel Act Mobility Program (aka IPAs), could also be used to recruit experts temporarily to provide surge support, expertise, and training to permanent staff/teams.

Section 3 - Data Inventory

Common Themes

Respondents to Section 3 of the RFI provided suggestions regarding data accessibility and reusability outside of the government. The common themes were:

- Prioritizing accessibility through machine-readable formats.
- Focusing on data value and reusability.

Key Feedback

Respondents stated that specific information is challenging to find due to the large volume of federal data. To remedy this, respondents suggested Federal Agencies control data growth to grow value and use it as a product across consumers.

Respondents would like the Federal Government to continue to invest in resources like data.gov, rectify datasets not found in any inventory, and extract data from PDFs into machine-readable formats. Respondents highlighted that metadata creation should be prioritized, as it is fundamental to working with publicly available data. Respondents also had specific recommendations for the General Services Administration to make additional investments to consolidate the Integrated Award Environment (IAE).

Respondents suggested agencies improve data discoverability in the IAE by:

- Cross-linking (e.g., Contract/Spend; streamlining number of unique identifiers).
- Incorporating more data dictionaries or 'glossaries' (e.g., USAspending.gov).
- Scanning and editing all federal websites referencing www.beta.sam.gov instead of www.sam.gov.

According to respondents, ways for agencies to make data more actionable include creating a classification for data elements and metadata, documenting data catalog data, exposing metadata as widely as possible, unlocking historical data or data trapped in documents, and layering search capabilities on top of data inventory processes. Respondents also noted that agencies should use a taxonomy with embedded metadata associated with each reported datum to report and understand the information collected easily.

Respondents suggested CDOs prioritize efforts to inventory and publish a catalog of high-value data. They emphasized that agencies should harmonize data inventory efforts across the government space and seek direction on meeting existing requirements.

Respondents noted that agencies should delineate dynamic data versus stagnant data and provide context for data as an inventory of use cases and examples alongside publicly available data. Additionally, clarification of types of data may help users understand what data means and how to use it.

Respondents emphasized that agencies should link to the source data for any tables included in their reports. Providing source data allows consumers to consider alternative analyses.

Section 4 - Data Sharing

Common Themes

Respondents to Section 4 of the RFI provided several suggestions for improving federal data-sharing practices. Common themes included:

- Using best practices for sharing programmatic data, implementing privacy-enhancing techniques, and creating/adopting government-wide policies.
- Implementing standards for data sharing.

Key Feedback

Respondents proposed Federal agencies consider modernizing their data management architectures by adding Data Virtualization to provide a Logical Data Fabric/Logical Data Warehouse. Modernization would allow for faster data sharing across agencies and the public and avoid costs associated with replicating, managing, storing, and securing data. Respondents identified Data Lakehouse architecture as a low-cost, flexible option for federal agencies to unify data, analytics, and AI workloads.

To improve privacy, respondents noted that agencies should consider creating a Logical Data Fabric to centrally manage security and governance across all data sources. Through this approach, policies enforcing privacy-protecting identifiers are built once, maintained centrally, and implemented for all data consumers down to the field level.

Respondents suggested the Federal Government encourage the creation of streamlined data-sharing processes and open-source identifiers, such as Legal Entity Identifiers, across agencies. Respondents also suggested privacy protections be prioritized and put in place through differential privacy and technical authentication methods.

Respondents flagged Privacy Preserving Record Linkages as an option for federal agencies to optimize and enhance current data matching efforts because these linkages would provide deidentification technology for PII, allowing data to be linked while maintaining privacy

To improve privacy, respondents noted that the Federal Government should consider creating a Logical Data Fabric to centrally manage security and governance across all data sources. Through this approach, policies enforcing privacy-protecting identifiers are built once, maintained centrally, and implemented for all data consumers down to the field level.

Respondents suggested that Federal agencies implement data sharing initiatives that focus on sharing data related to a 'customer journey.' They explained that automation can help map data and recommendation systems should proactively suggest data or user groups for sharing. For example, analytics-first system architectures allow data sharing on top of any public cloud infrastructure, decoupling data usage efforts from storage location.

To encourage data sharing and integration, respondents recommended the Federal Government direct legal counsels to work with the CDO Council to help clarify and evaluate federal laws associated with data sharing and consider engaging with Congress in support of a National Secure Data Service.

Respondents suggested the Federal Government implement data standards and expand metadata schema emphasizing privacy and interoperability while prioritizing publicly accessible data

Section 5 - Value and Maturity

Common Themes

Respondents to Section 5 of the RFI provided several suggestions for improving the value and maturity of federal data. Common themes included:

- using different methods to measure the value of data.
- increasing and improving communication methods associated with data.
- referencing examples and guidance in conducting maturity assessments.

Key Feedback

Respondents encouraged the Federal Government to consider different approaches to data valuation. Some existing approaches to data valuation are market-based models (which calculate data's utility in terms of cost and revenue/profit), economic models (which estimate data's utility in terms of economic and public benefit), and dimensional models (which extend the above models to estimate utility based on categories or dimensions—both data-specific and contextual.)

Respondents recommended the Federal Government consider measuring the value of federal data through scorecards summarizing data usage. The Forest Service's "High Conservation Value" approach was offered as an example of a data valuation approach for developing metrics (diversity, community needs, cultural values, ecosystem services, etc.).

Respondents highlighted that CDOs should utilize diverse communication channels to increase the value and use of data for stakeholders and to communicate progress on and values of data governance efforts. Examples provided include newsletters, an implementation website (e.g., SharePoint site), signage, lunch and learns, town hall forums, informational videos, and through other established meetings.

To gauge existing data capacity and needs accurately, respondents suggested CDOs repeat maturity assessments annually/regularly. They emphasized that these maturity assessments should be performed consistently and continuously across the Federal Government for purposes beyond compliance. The FISMA scorecard is an example of this.

Respondents identified the CMMI Data Management Maturity model as an example of a straightforward, refined, and familiar framework for maturation for data management capabilities

Section 6 - Ethics and Equity

Common Themes

Respondents to Section 6 of the RFI provided several suggestions for improving and monitoring federal data ethics practices. Common themes included:

- Building out and clarifying guidance related to ethics and equity.
- Implementing methods to mitigate bias and to ensure that federal data ethics strategies are sufficient.
- Taking measures to increase the transparency, availability, and sharing of best practices associated with ethical data use.

Key Feedback

Respondents suggested the CDO Council provide clear, unified guidance regarding ethics and equity standards for data using the Federal Data Strategy and other existing frameworks. They also suggested the Council collaborate with ethics-focused organizations outside of government to encourage the application of best practices and continuous improvement in developing this guidance.

Respondents flagged that the Data Ethics Framework lacks specific guidance on collecting race, ethnicity, gender, and disability data. They recommended the CDO Council expand the framework to address data collection without causing harm and address responsible use of data to serve vulnerable communities better.

Respondents recommended the CDO Council and the Interagency Council on Statistical Policy create a permanent data ethics working group within the Council to ensure the Data Ethics Framework continuously meets emerging needs, provide resources and guidance to agencies, and partner with relevant professional associations for ongoing education and training on data ethics.

Respondents stated that the CDO Council should advocate that agencies publish more statistical data related to race, ethnicity, demographics, and other diversity measures. They also emphasized that the Data Ethics Framework should push for increased publication in a central repository with metadata and metrics on completeness, timeliness, and quality. Finally, respondents stated that the Framework should also create a secure environment for partners to share this kind of data to assist front-line workers.

Respondents noted that the CDO Council should examine the entire lifecycle of data governance through commercial software and industry tools to maximize the intersection of data ethics with Diversity, Equity, Inclusion, and Access and to ensure that the technology used meets appropriate standards.

Respondents recommended the CDO Council issue ethics and equity scorecards of how tenets are applied in order to increase government data transparency and availability and help establish a baseline and identify potential biases and inequities in analytics efforts and government services.

To keep improving trust and transparency in federal data, respondents requested relevant agency contact information be publicly available and those individuals be responsive to feedback from and partnerships with external organizations.

Section 7 - Technology

Common Themes

Respondents to Section 7 of the RFI provided several comments regarding trends in data technology. Common themes included:

- Partnering with Chief Information Officers and other stakeholders to leverage technologies.
- Promoting consistency across agencies.
- Focusing on advances in cloud data management and privacy and protection.

Respondents also provided examples of frameworks to be used in evaluating data infrastructure.

Key Feedback

When adopting new technology to help support data functions, respondents encouraged the CDO Council is to partner with Chief Information Officers (CIOs) and other relevant stakeholders to leverage technologies, where possible, to avoid “reinventing the wheel” and to support and accelerate proactive approaches to ensuring ethical, legal, and policy appropriate discovery, use, and sharing of data and information.

Respondents stated that CDOs, CIOs, and Chief Information Security officers need to work together to develop zero trust data security strategies and meet the objectives outlined in the Federal Zero Trust Architecture Strategy. They emphasized that this effort would ensure effective and efficient solution delivery and avoid data management and analytics challenges.

Respondents recommended the Federal Government build partnerships between CDOs and Chief Experience Officers as technology provides more opportunities than ever to take a “data-first” view and close the loop on engagement with customers and citizens.

Respondents noted that agencies should work with the computer science, mathematics, cybersecurity, privacy, and statistical communities to explore and develop risk-based policy frameworks and approaches to support enhanced open data, data sharing, and linking datasets while ensuring protection for sensitive or personally identifiable information.

Data Strategy, Data Cataloging, and a Data Platform approach were highlighted by respondents as “tools in the toolbox” for Federal CDOs to help drive full implementation of the Federal Data Strategy.

Respondents suggested the Federal Government drive better agency delivery by advising every agency to provide data capabilities in the cloud, leveraging managed services in the cloud mandating both SQL access and/or Application Programming Interface access to agency data, and, based on mission needs, using Artificial Intelligence/Machine Learning to deliver outcomes.

Respondents emphasized that blockchain is gaining ground as a decentralized authoritative source, and suggested the Federal Government consider this technology as it moves towards citizen-centric data ownership.

Respondents suggested that the CDO Council build capabilities to integrate and support differential privacy, digital watermarking, federated learning, homomorphic encryption, and other secure data transfer and privacy protection technologies. Respondents also noted that it is important for CDO's to

facilitate interactions within their agencies with key stakeholders to accomplish the strategies, as opposed to over burdening their offices with ownership responsibilities.

Respondents provided information about the Enterprise Data Management Council's (EDMC) Data Capabilities Assessment Model and suggested that agencies use this model to evaluate their data infrastructures. Respondents also noted that agencies can use EDMC's Cloud Data Management Capabilities Framework to determine best practices for managing data in the cloud.

Respondents recommended that the CDO Council consider approaches published by other countries. For example, the European Data Strategy aims for a single data market, allowing data to flow freely across the EU.

If you have questions or would like more information about the case studies, contact cdocstaff@gsa.gov.



www.cdo.gov | cdocstaff@gsa.gov