

ATTACHMENT I – PROJECT TOPIC

Data Usage Platform as a Federal Data Asset

Key Objective

The objective of this project is to research and develop a robust and sustainable framework that will enable the federal data ecosystem to better understand the uses of federal data in support of a potential, future National Secure Data Service (NSDS). Related work includes the dashboard prototype efforts of the Democratizing Data Initiative, which promotes the use of federal data and assets for evidence building (<https://democratizingdata.ai/>). This project will produce possibilities for a future, state-of-the-art, updatable publicly accessible platform that provides information on federal data usage as part of the National Secure Data Service demonstration project.

Background

The Democratizing Data Initiative is a cross-agency, multidisciplinary effort to demonstrate the value of statistical data through aggregated usage statistics that are displayed on dashboards and for use in other tools. The initiative was a response to the requirements for open data as part of the Foundations for Evidence-based Policymaking Act of 2018 and other federal transparency initiatives. Understanding the uses of federal statistical data can support collaboration across various research and policy communities and promote cross-disciplinary approaches to research and evidence-based policy making. Highlighting use and application of federal data assets can also support statistical agencies' business needs and decision making related to stakeholder requirements, stakeholder engagement, and supporting investments in data programs and other efforts. The Advisory Committee on Data for Evidence Building recommended the development and application of usage statistics to respond to the future needs of a National Secure Data Service (<https://www.bea.gov/system/files/2022-10/acdeb-year-2-report.pdf>). Building on prototype efforts, this project aims to explore alternative dashboard technologies, approaches, techniques, and methods to meet the sustainability and transparency needs of the federal statistical system and to inform the efforts of the National Secure Data Service Demonstration Project as required by the 2022 CHIPS and Science Act.

Information Gaps

This project will answer the following questions:

- What is user feedback on the Democratizing Data pilot dashboards and tools (found here <https://democratizingdata.ai/>)? What information do users report is useful in the current dashboard and what additional information would be useful? User groups may include but are not limited to researchers, analysts, policymakers, federal agencies, program managers or data journalists.
- What additional types of federal data, features and functions of the dashboard or other interface are required to support evidence-building use cases for a wide variety of stakeholders, federal agencies, and consumers of federal statistics? Does this vary by sector, geographies, federal agencies, and user groups?

- How can these usage data be leveraged in a profile to inform tiered data access needs, transparency priorities, and open data requirements for federal agencies?
- What open data software, data science, best practices, and other cutting-edge technologies can be used to generate aggregated usage data that meet the quality and sustainability requirements of federal agencies?
- What are the resources required for flexibility, sustainability and scalability of the dashboard platform and related tools?
- Do these usage statistics offer a vision for understanding the health and progress of the federal statistical system?
- How can these usage data be used as a shared service to inform public trust of official statistics, and engagement in a possible, future National Secure Data Service (NSDS)?

Objectives

The objectives for the Democratizing Data Platform as a Federal Data Asset project include:

1. Identify sustainable, state of the art, updatable solutions, tools, or frameworks to ingest, aggregate and deploy usage statistics. Approaches may include but are not limited to data science techniques and tools such as machine learning and natural language processing.
2. Identify, interview, and document use cases across sectors and subgroups to ensure usability and accessibility of usage data for collaboration, evidence building, and other purposes.
3. Identify alternative data, formats, and tools (e.g. APIs) to present and complement usage data.
4. Develop, refine, and/or enhance dashboard interface, functions, and features to encourage use across various stakeholder communities.
5. Identify workflow processes and other mechanisms to ensure efficient, quality, timely and up-to-date usage data, and information.
6. Implement technologies, techniques, and processes to build a prototype dashboard platform and tools using publicly available federal data and information.
7. Identify approaches and best practices to implement data quality flags to communicate reliability and fitness for use of usage statistics.
8. Building on findings and lessons learned, advise on the integration of the platform and tools within the federal data ecosystem and possible, future NSDS.

Deliverables

At a minimum, offerors will provide the following if selected for award. Additional deliverables may be required.

- Use case report and gap analysis based on interviews, focus groups or other data collection methods.
- Plan for development of data usage dashboard and tools.
- A prototype data usage dashboard or other presentation and tools.
- Usage, specification, and code documentation to accompany the prototype data usage dashboard and tools.
- Communication Plan for implementation and deployment across federal agencies.
- Training data (if applicable), and any other data created under this award.