

May 31, 2022

RE: National Center for Education Research, Institute of Education Sciences on the Use of Large Datasets

To Whom it May Concern,

The Data Foundation is a non-profit organization that seeks to improve government and society by using data to inform public policymaking. Our Data Coalition Initiative is America's premier voice on data policy, advocating for responsible policies to make government data high-quality, accessible, and usable. An important part of that effort is supporting the development and implementation of the Foundations for Evidence-Based Policymaking Act (Evidence Act), which aims to improve the efficiency and effectiveness of government services through generation and use of data and evidence.

In line with the Evidence Act goals, the creation, maintenance, sharing, and use of large datasets can allow both government and non-government researchers access to information that informs the development and analysis of policies that can improve government services. The Data Coalition's expertise on federal data practice and policy lends itself to inform the National Center for Education Research's (NCER) request for information on the existence and use of large datasets to address education research questions, particularly Question #8.

#8. What are the best practices for creating new datasets or linking existing datasets and sharing them with researchers (open or restricted use) while prioritizing the privacy of individuals and adhering to local, State, and Federal laws? What barriers and limitations exist?

While we know the federal government collects vast amounts of data that have the potential to contribute to new insights in education, among other areas, the complexity of data sharing is severely limited in the current system. Data linkage and sharing across government and with researchers occurs through a variety of different data sharing agreements that exist at different agencies and across different levels of government, creating a complex and inaccessible system.

One way that the federal government can link and share datasets is through a national data service. A data service would provide a single, central location where government data can be temporarily linked and shared while maintaining strong privacy protections. The Data Coalition supports the National Secure Data Service Act, which builds on the bipartisan and unanimous recommendations from the U.S. [Commission on Evidence-Based Policymaking](#) from 2017, a consensus proposal from the [National Academies of Sciences, Engineering and Medicine](#) in 2017, and a suggested roadmap published by the [Data Foundation](#) in 2020. The proposed legislation creates an expectation for the National Science Foundation (NSF) to make rapid progress in launching a data service and transparently supporting government-wide evidence-building activities.

Under the proposed legislation, the data service at NSF must adhere to federal privacy laws, including the Confidential Information Protection and Statistical Efficiency Act of 2018 (CIPSEA). This law was reauthorized by Congress with bipartisan approval in 2018, establishing one of the strongest government privacy laws in the world, including strong criminal and civil penalties for misuse. Certain federal statistical agencies are currently authorized under CIPSEA and can share, access, and use data from across government. However, agencies that are not currently CIPSEA-designated agencies, such as NCER, cannot access CIPSEA-protected data. A secure data service would allow non-CIPSEA agencies access to CIPSEA data, facilitating new insights across research areas.

An infrastructure such as the National Secure Data Service would provide a resource for federal agencies, researchers, and data analysts to responsibly produce insights that can address potential research questions. The linkage capacity and incorporation of strong privacy and confidentiality mechanisms can be adapted to service research centers, such as NCER, across the federal government to garner evidence that can improve agency effectiveness and better work to achieve their missions.

Thank you for the opportunity to provide comment on this very important issue, and we hope to continue to support your efforts in the successful use of data to ensure the federal government is providing high-quality services to its taxpayers. Please contact me at corinna@datafoundation.org if you have any questions or would like to discuss the Data Coalition's interest in this matter further.

Sincerely,

Corinna Turbes
Managing Director
Data Coalition