
A Blueprint for Implementing the National Secure Data Service:

Initial Governance and Administrative Priorities for the National Science Foundation

June 2022



Authors



Nancy Potok, Ph.D.

Former Chief Statistician of the United States
White House Office of Management and Budget
Data Foundation Board Member



Nick Hart, Ph.D.

President & CEO
Data Foundation

Acknowledgments

The Data Foundation and authors thank the Schmidt Futures for its generous support for this project. The authors also thanks the independent reviewers who provided constructive comments on earlier drafts and the members of the Advisory Committee on Data for Evidence Building who convened public discussions that contributed to the ideas in this report.

Disclaimer

This paper is a product of the Data Foundation, with funding support provided by the Schmidt Futures. The findings and conclusions expressed by the authors do not necessarily reflect the views or opinions of the Data Foundation, its funders and sponsors, or its board of directors.

Disclosure

Both authors of this paper were affiliated with the U.S. Commission on Evidence-Based Policymaking. Dr. Potok was an appointed member of the Evidence Commission and Dr. Hart served as the policy and research director.

Tables of Contents

- 4 Executive Summary**
- 7 Introduction to the National Secure Data Service Concept**
- 10 Essential Attributes and Functions of a National Secure Data Service**
- 12 Potential Governance Models for the NSDS**
- 18 Blueprint for NSDS Governance Model**
- 25 Summary Findings**
- 26 Recommendations on Next Steps**
- 27 Resources**



Executive Summary

The concept of the National Secure Data Service (NSDS) has gained substantial momentum and enthusiasm since it was first proposed by the U.S. Commission on Evidence-Based Policymaking (Evidence Commission) in its 2017 Report to Congress and the President.¹ Since 2017, Congress passed the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), which established the Advisory Committee on Data for Evidence Building (ACDEB) to explore questions and implementation options for the NSDS.² The ACDEB is charged with providing recommendations to the Director of the White House’s Office of Management and Budget (OMB) on “...how to promote the use of federal data for evidence building with a particular emphasis on evaluating the need for and value of a National Secure Data Service.”³ The ACDEB launched in October 2020 to begin considering questions related to the NSDS.⁴

In 2020, the Data Foundation published *Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building*, which presented a broad vision for how to align the NSDS within the framework offered by the Evidence Commission and the Evidence Act. That vision is to build evidence that supports public policymaking in a collaborative ecosystem, including but extending beyond federal statistical agencies.⁵ The paper specifically examined four options for an overall structure and location for the NSDS, recommending that it operate as a specific type of government-owned and contractor-operated facility to accommodate public-private partnership. The paper also identified key functions of the NSDS and recommended that it be housed at the National Science Foundation (NSF) and contracted out through the National Center for Science and Engineering Statistics (NCSES), one of the 13 designated federal statistical agencies.

One major limitation to launching the NSDS has been that while there have been high-level discussions, including from the Data Foundation, about what the NSDS can do and where it should be situated, there has been relatively little presented publicly about an operational strategy for implementation at NSF. This report seeks to address this gap by providing a blueprint to begin NSDS implementation, including structure, governance, and activities in several key areas building off the *Modernizing U.S. Data Infrastructure* report. Standing up and maintaining an NSDS successfully will require substantial support, technical capability, governance, and oversight.

The core activities of the NSDS include hosting a secure infrastructure where government and non-government researchers and program evaluators would: (1) submit proposed projects for approval; (2) link and access data for research and analyses; and (3) have project results privacy-protected, then prepared for public dissemination. However, in considering implementation options, it is essential that **the initial structure and functions of the NSDS serve the needs of a core NSDS and be flexible enough to add additional functions if they are viewed as valuable as the NSDS matures.** Further functionality could be added through establishing or designating specialized research nodes that become partners in the NSDS network. There is no recommended timeframe for when the NSDS can be considered mature—if successful, it should be demonstrating value and continuous learning that gains a consensus among its users that adding functionality is beneficial. In other words, the dynamic nature of the NSDS must be planned for at the outset in the program design.

This report examines three existing models in assessing the best fit for the NSDS across multiple dimensions such as a concept of operations, governance, oversight, and funding, while also building on the Data Foundation’s *Modernizing U.S. Data Infrastructure* report from 2020.⁶ The three models we examined are:

- Government-Owned, Government-Operated (GOGO)
- Government-Owned, Contractor-Operated (GOCO)
- Government-Owned, Grantee Institution-Operated (GOGIO)

This report proposes a combined structure incorporating elements from the three models that would best support the attributes of the NSDS under a government-owned, contractor-operated facility infrastructure model, as proposed in the Data Foundation's *Modernizing U.S. Data Infrastructure*. It discusses how the NSDS might get started around its core functions to enable later expansion, but most importantly, the report offers a blueprint for how NSF can begin this work quickly, efficiently, and transparently. We have extended the ACDEB framework published in 2021 to include the concept that the NSDS should also be part of a network, in which it plays the central role of coordinating a group of nodes which make up a collaborative system guided by a coherent governance process. Serving as a hub is fundamental to the success of the NSDS beyond the demonstration phase and should include integration in the federal statistical system and partnerships with administrative program offices, state and local governments, tribal governments, and the research community.

Findings

Below are our findings and recommendations on next steps for consideration of NSF, NCSES, and the ACDEB, which provide a blueprint for launching the NSDS as NSF.

- The initial structure and functions of the NSDS, while it is a pilot project, should put a foundation in place that would both continue to serve the needs of a core NSDS and be flexible enough to add additional functions if they are viewed as valuable as the NSDS matures.
- There is no recommended timeframe for when the NSDS can be considered mature—if successful, it should be demonstrating value and continuous learning that gains a consensus among its users that adding functionality is beneficial.
- The NSDS should adopt a hybrid governance structure incorporating the attributes that would best support the functions of the NSDS under the infrastructure model proposed in the first paper, that is, a government-owned contractor-operated facility.
- Under a government-owned, contractor-operated model, there would be specific roles for the various aspects of the governance structure, including for NSF, NCSES, a Policy Steering Committee, the Contractor Entity, a Board of Governors, a Research and Technology Advisory Board, NSDS Operations, Collaborative Partnerships, Client Stakeholders, a Project Approval Committee, and a Project Appeals Board.
- Efficient, transparent, and timely business processes are essential elements for a successful NSDS. The NSDS would be an ideal place to host the application and approval process, provide core services, and help facilitate locating research at other facilities if that would provide a clear benefit. However, the NSDS management will need to identify where agency policies are creating bottlenecks and assist agencies in finding more efficient approaches to meeting their requirements.
- Sustainable funding is imperative for the success of the NSDS. The base funding should come from a federal appropriation to NCSES, which would then be used to fund a contract with a contractor entity and to pay for program management at NCSES. Funding should also come from user charges, which should be easily understandable and published on an equitable fee-for-service schedule based on usage that does not constitute a barrier to access for researchers. The contractor should be able to also collect revenue as part of partnerships and cooperative agreements.

Recommendations on Next Steps

- NCSES should begin to structure the competitive bidding for an NSDS pilot so that a Request for Proposals could be released shortly after enactment of authorizing legislation and NCSES receipt of appropriations. NCSES should use America's Data Hub where appropriate to move forward portions of the NSDS concept, such as coordination of data linking research centers that currently exist.
- While the competitive bidding process is underway, NCSES should work with OMB to establish the Policy Steering Committee and other governance structures.
- The Interagency Council on Statistical Policy (ICSP) should seriously consider structuring the substantial work that has already been done to implement the statistical agency pilot application portal to adapt it to the NSDS application approval process.

There are many questions that will remain left unaddressed about how to conceptualize, establish, launch, staff, operate, interact with, and manage the NSDS at NSF. We did not set out to address every conceivable question in this blueprint, but rather to identify a starting point for NSF on the major questions of governance and operations. We intend for this to be a productive starting point for NSF and to inform the deliberative process of the ACDEB members. Our findings, building from our prior research, suggest a potential path that is reasonable, and a strong improvement from the status quo.

While we offer this blueprint as a starting point, the ultimate success of the National Secure Data Service rests with the data and evidence community. The statistical system, the evidence community, and the country will all be better off with the NSDS than without it. Advancing implementation of the National Secure Data Service is a safe, practical, and necessary way to ensure our country's leaders will have the data and evidence they need to make important decisions. The blueprint we offer in this report for NSF proposes a practical strategy for prioritizing how NSF could begin to implement the NSDS in the near-term. Echoing the sentiment of the Evidence Commission – it's time for the vision to become reality.

Introduction to the National Secure Data Service Concept

The concept of the National Secure Data Service (NSDS) has gained substantial momentum and enthusiasm since it was first proposed by the U.S. Commission on Evidence-Based Policy (Evidence Commission) in its 2017 report to Congress.⁷ The Evidence Commission's recommendation to establish a secure data service has garnered attention from stakeholders, researchers, and government officials as a viable and realistic solution for addressing data linkage challenges that exist in our society today.

The NSDS is a concept that complemented work undertaken by the National Academy of Sciences Committee on National Statistics (CNSTAT), which explored options for increasing data sharing among federal statistical agencies. The focus of CNSTAT was modernizing and improving federal statistics, while the Evidence Commission had a broader mandate from Congress to look at how capacity for evidence-building across the federal government could be improved and sustained. In its final report to Congress and the President, the Evidence Commission declared that it envisioned "a future in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy."⁸

The Evidence Commission's central recommendation was to establish an NSDS to facilitate data access for researchers both inside and outside government for evidence-building activities. The Commission also stated that using an NSDS could enhance privacy protections for sensitive data through use of technological advances, even while enabling linkage of various data sets from different sources to create better information on which to make public policy decisions. However, the Evidence Commission did not provide detail on the specifics of design, governance, implementation, funding, or operations of such a government-wide service. While some of these activities were discussed by the commissioners and are documented in deliberative materials that are now public, the final recommendations of the Commission provided light guidance on many design features of an NSDS.⁹

Congress subsequently passed the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act), which included 11 recommendations from the Evidence Commission.¹⁰ The Evidence Act did not take explicit action on the NSDS but signaled interest in continuing to explore the concept by establishing the Advisory Committee on Data and Evidence Building (ACDEB).¹¹

The ACDEB is charged with providing recommendations to the Director of the White House's Office of Management and Budget (OMB) on "...how to promote the use of federal data for evidence building with a particular emphasis on evaluating the need for and value of a National Secure Data Service."¹² The ACDEB launched in October 2020 to begin considering questions related to the NSDS and supporting implementation of the Evidence Act.¹³

Also during 2020, the Data Foundation published *Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building*, which presented a broader vision for how to align the NSDS within the framework offered by the Evidence Commission and the Evidence Act.¹⁴ The paper presented the joint goals of the Evidence Commission and Evidence Act that building evidence to support public policymaking in a collaborative ecosystem included but extended beyond federal statistical agencies. The paper specifically examined four options for an overall structure and location for an NSDS, recommending that it operate as a special type of government-owned and contractor-operated facility to accommodate public-private partnerships. The paper also identified key functions of the NSDS and recommended that it be housed at the National Science Foundation (NSF) and contracted out through the National Center for Science and Engineering Statistics (NCSES), one of the 13 designated federal statistical agencies.

The Data Foundation's paper, *Modernizing U.S. Data Infrastructure*, was intended to inform the work of the ACDEB. It specifically offered a series of principles and questions for consideration by the ACDEB that would benefit from the wide range of expertise and perspectives gathered in a federal advisory committee process. The paper was provided during the early stages of the committee's formation to all of the members for background and was presented and discussed during a public session in April 2021.¹⁵

In July 2021, shortly after the ACDEB discussions about the Data Foundation's paper, the U.S. House of Representatives passed the National Secure Data Service Act, a proposal that effectively adopted the Evidence Commission's recommendation to establish an NSDS with the Data Foundation's proposed approach, aligning the capability with the NSF.¹⁶ The legislation passed the House in January 2022.¹⁷ Although it was a slight variation from the original Evidence Commission proposal, the Evidence Commission's co-chairs publicly endorsed the legislation as a practical and responsible strategy for adopting their recommendations.¹⁸

BOX 1: What is the National Secure Data Service Act? ¹⁹

The federal government's data infrastructure is largely decentralized. Individual agencies and programs may collect data without sharing or using information already collected by other parts of government. This imposes undue burdens on the American public and businesses through repeated reporting of information the government already has. Creating a capacity to securely share information while protecting confidentiality and deploying other privacy safeguards offers tremendous potential for developing new insights and knowledge to support statistical analysis and summary-level information relevant for evidence-based policymaking and practice.

The NSDS Act calls for the establishment of a demonstration project "to develop, refine, and test models to inform the full implementation of the Commission on Evidence-Based Policymaking recommendation for a government-wide data linkage and access infrastructure for statistical activities conducted for statistical purposes."²⁰ The demonstration project is to be housed at NSF and operated "directly or via a contract that is managed by the National Center for Science and Engineering Statistics."²¹

The data service is designed to supplement rather than displace any existing, successful, and sufficiently secure data linkage arrangements. Statistical agencies engaged in production-level data collection, sharing, and publication for the development of federal statistical indicators will receive additional capabilities from the National Secure Data Service but could retain existing practice.

The legislative interest expressed by Congress in the NSDS—which was built on the 2020 report from the Data Foundation, the Evidence Commission, and past work from CNSTAT—is the motivation for this report. To support the ongoing policy dialogues about effective and timely implementation of the NSDS, including to inform the ACDEB, the authors of the 2020 Data Foundation publication provide extended, renewed, and refreshed analysis to address the most critical initial starting points for NSDS implementation at NSF. This is not to say that the issues and topics discussed in this report will not merit further discussion, including among ACDEB members, the Interagency Council on Statistical Policy (ICSP), OMB, and other consultative bodies. However, given the import and impact of our initial 2020 report on the current processes, we thought it would be useful to provide additional perspectives and a proposed blueprint to inform the next steps for implementation. We also seek to address some of the questions that we left unanswered in our 2000 report.

We recognize that the ACDEB continues its thoughtful and deliberative process, building on a set of recommendations from October 2021.²² The ACDEB stated in its interim report that it had a "focus on expanding access to non-public data for evidence building, and that the evidence-building ecosystem would both benefit and support the success of an NSDS."²³ The ACDEB report also identified key functions for the NSDS and areas for potential exploration as it develops its final report, expected in October 2022. This report can hopefully contribute to and support the ACDEB considerations in 2022.

Concurrent with these advisory board and legislative activities, the NCSES at NSF established a pilot project called America's Data Hub that envisions to "...coalesce the statistical community and involve the collection, interpretation, analysis, and dissemination of data on science, engineering, technology, and research and development, including innovative services that involve data access and linkage and that address security and privacy concerns."²⁴ NSF used its unique "Other Arrangement" contractual authority to establish the hub, in order to provide the flexibility to bring together stakeholders.²⁵ Under this special authority, NSF may "...initiate and support basic scientific research and programs to strengthen scientific research potential...by making contracts or other arrangements (including grants, loans, and other forms of assistance)."²⁶

According to NSF, the purpose of the Data Hub is to pilot innovative technologies and processes for a data service that supports government-wide evidence-building activities. The hub is an open consortium composed of members who enroll online. Proposals for projects are floated to the membership, who then submit proposed solutions. The solutions are evaluated and selected by the Technical Advisory Board, which consists of elected members of the consortium. The hub is not an NSDS, which by contrast is envisioned as a place where large numbers of researchers and practitioners both inside and outside government can access confidential, protected linked data provided by federal agencies for the researchers' own statistical activities (including program evaluations in an agency's learning agenda).²⁷ While the current structure of America's Data Hub is significantly different than that envisioned for the NSDS, it potentially presents an opportunity for alignment of core concepts and ideas.

One major limitation to launching the NSDS has been that while there have been high-level discussions, including from the Data Foundation, about what the NSDS can do and where it should be situated, there has been relatively little presented publicly about an operational strategy for implementation of NSF. This report seeks to address this gap by providing a specific blueprint for NSDS implementation, including structure, governance, and activities in several key areas building off the *Modernizing U.S. Data Infrastructure* report published in 2020. We continue to recognize and support that the core activities of an NSDS include hosting a secure infrastructure where researchers and program evaluators would: (1) submit proposed projects for approval; (2) link and access data for research and analyses; and (3) have project results privacy-protected and prepared for public dissemination. Standing up and maintaining an NSDS successfully will require substantial support, technical capability, governance, and oversight.

NSF will need to make additional decisions about its ability to provide various functions. In addition, NSF will need to consider the value of additional functions, which could potentially include:

1. serving as a central point of coordination and information sharing for the decentralized research centers around the U.S. that currently host sensitive federal data or support researchers engaged in evidence-building;
2. having the NSDS host a statistical agency research project application portal and approval process and opening it to non-statistical federal agencies;
3. serving as a hub for research in data science methodology to enhance privacy and confidentiality protections;
4. serving as a clearinghouse for best practices on privacy protection and data linkage;
5. developing (or at least providing) metrics on assessing the quality of combined data; and
6. providing useful information to researchers about other similar research projects and published articles in order to foster increased collaboration around topics of public policy interest.

This is understandably a lot to expect for a demonstration project in government, which also means there is a need to ensure alignment across core constituencies in government and stakeholder communities to support efficient, timely implementation. This report discusses options for how the NSDS might get started around its core functions to enable later expansion, but most importantly it offers a blueprint for how NSF can begin this work quickly, efficiently, and transparently.

Like the Data Foundation's *Modernizing U.S. Data Infrastructure* report, the intent of this report is not to duplicate the efforts of the ACDEB or other advisory bodies, but rather to extend our earlier work and to inform the ACDEB's Year-2 final report.²⁸ We look forward to the ACDEB's advice to the OMB Director on the design of the NSDS and other emerging issues as part of its final report due in October 2022. Implementation will be difficult for the NSDS and the stronger the collective advice, the better NSF and supporting agencies will be able to establish an effective system that successfully accomplishes the important evidence-building goals and priorities.

Essential Attributes and Functions of a National Secure Data Service

The Data Foundation's *Modernizing U.S. Data Infrastructure* report explored multiple approaches for implementing a National Secure Data Service, including: (1) establishing a new agency at the Commerce Department; (2) re-tasking an existing agency at the Commerce Department; (3) creating a new Federally-Funded Research and Development Center (FFRDC) at the National Science Foundation, a type of government-owned, contractor-operated partnership model; and (4) launching a public-private partnership in a university consortium. After reviewing the capabilities for each option, we recommended that a new center be launched at NSF as a public-private partnership using a government-owned contractor-operated model. The report also outlined the necessary attributes of a data service. Similar attributes had been described by the CNSTAT panel, the Evidence Commission, and the ACDEB Year-1 report.²⁹

For the sake of consistency and ease, this report adopts the wording of the attributes described in the ACDEB Year-1 Report as its basis of discussion and analysis. Box 2 contains short definitions of each of the attributes.

BOX 2: Attributes of a National Secure Data Service

- **Transparency and Trust:** Ability to (1) explain clearly what activities are being undertaken and their benefits, (2) be guided by a set of standards for maintaining the public trust, and (3) include a representative oversight structure.
- **Independence:** Empowerment to set strategic priorities, operate apart from policy and political offices, support objective analyses, and protect privacy and confidentiality at the same level as a federal statistical agency.
- **Scalable Functionality:** Ability to recruit, hire, and retain skilled personnel who understand data curation record linkage, machine learning, statistical computing, data transmission, data encryption, the legal and regulatory framework around various data sets, identity disclosure avoidance, and IT architecture and security that can be scaled up to meet needs. The ability to put in place and maintain efficient business processes that do not become a bottleneck for researchers or bureaucratic barriers for staff.
- **Sustainability:** Support from a secure funding source and secure placement with a federal agency, with sufficient funding and staffing to stand up and grow the services.
- **Oversight and Accountability:** Mechanisms are in place to assure responsible, effective, ethical, and legal compliance with its charge and mission.
- **Intergovernmental Support:** Capability to coordinate and collaborate across government entities beyond the federal government, including key data providers and partners.
- **Legal Authority to Protect Privacy and Confidentiality:** Authority commensurate with Title 3 of the Evidence Act, known as Confidential Information Protection and Statistical Efficiency Act (CIPSEA), as well as the Privacy Act, along with the ability to develop and deploy new privacy preserving technologies.
- **Legal Authority to Collect Data from Federal Agencies:** Authority consistent with capabilities provided to statistical agencies under CIPSEA Part D, that is, a presumption that federal data should be made available to statistical agencies for evidence-building purposes unless expressly prohibited by statute.
- **Accessibility:** Embracing of plain language, diverse stakeholder engagement in governance and ongoing assessments of timeliness in data availability, access, and use, which includes promotion of equitable access through data access processes and procedures and a clear and cost-neutral fee structure.

The attributes provide the basis for assessing different options and considerations for implementation and help frame or weigh choice points and solutions for implementation as will be discussed below. This is an important starting point for the discussion that follows because we agree in principle with the framework offered by the ACDEB in its Year-1 report. This includes not just the attributes, but also in recognizing that the NSDS is a construct that is a philosophy, a place, and a service.

Nevertheless, we extend the ACDEB framework with one additional leg—the **NSDS plays a central role in a network, responsible for coordinating a group of nodes that make up a collaborative system guided by a coherent body and governance process**. While NSF will have to make a determination about implementing this approach, serving as a hub is fundamental to the success of the NSDS beyond the demonstration project phase, including integration into the federal statistical system and building partnerships with administrative program offices, state and local governments, tribal governments, and the researcher community.

A well-functioning NSDS must be dynamic to meet emerging technological and systemic needs. For the NSDS to achieve such a dynamic nature, careful consideration must be given to the type of governance processes and procedures selected. In this report, governance is intended to be a broad term that is inclusive of the processes and structures for organizing guidance, administrative direction, establishing policies and procedures, providing oversight, facilitating accountability and transparency, and fostering engagement and feedback. There are, of course, other dimensions for consideration as well such as operations, funding, and technical capabilities, among others.

When considering implementation options, it is essential that **the initial structure and functions of the NSDS put a foundation in place that can serve the needs of a core NSDS and be flexible enough to add additional functions if they are viewed as valuable as the NSDS matures**. Further functionality could be added through establishing or designating specialized research nodes that become partners in an NSDS network. There is no recommended timeframe for when the NSDS can be considered mature—if successful, it should demonstrate value and continuous learning that gains a consensus among its users that adding functionality is beneficial. In other words, the dynamic nature of the NSDS must be planned for at the outset in the program design.

BOX 3: Out of Scope Operational Questions for Future Consideration

The concept of operations includes examining where the NSDS fits into the data ecosystem. That is, is it solely a service that links federal statistical data in a secure environment? If yes, why is it necessary or an improvement over the facilities that currently have the capability of linking data in a secure environment? Should it have a monopoly on these activities? How extensive is the NSDS's role as a coordinator for the decentralized evidence-building research ecosystem currently in place across the nation, providing a mechanism for coordination of best practices and building collaboration across important public policy areas? For example, should it set standards for quality or metadata or just provide a repository for information about the various standards being used?

Are there constraints on the types or sources of data for the NSDS. For example, will data from private sources be combined with traditional data from statistical agencies as approved projects? Would non-federal data be combined with other non-federal data, or would such linkages be deemed out of scope? These questions have implications for the design of the NSDS, including how state, local, tribal, and private data may be incorporated.

These key questions require substantial input from the stakeholder community, including data providers, owners, researchers, analysts, and disseminators, both inside and outside the federal government. Addressing these questions is beyond the scope of this paper, but they are all potential topics for the ACDEB or other advisory bodies to consider. Within the scope, however, is a recommendation for an infrastructure design that is flexible enough to accommodate these different concepts of operation as they are considered either during a pilot project or afterward.

Potential Governance Models for an NSDS

There are multiple approaches that could be considered reasonable and appropriate for establishing a governance structure for an NSDS based on approaches used in different parts of the government today. This report identifies and presents three existing models to assess the best fit for the NSDS across multiple dimensions such as a concept of operations, governance, oversight, and funding, and also builds on the Data Foundation's *Modernizing U.S. Data Infrastructure* report from 2020.³⁰

Considering the attributes, the ACDEB framework, and the collaborative nature of the NSDS, for this report the three models we reviewed would all enable the core functions of an NSDS mentioned earlier, including hosting a secure infrastructure where researchers could: (1) submit proposed research projects for approval; (2) link and access data for research and analyses; and (3) have research results privacy protected then prepared for public dissemination. The three models we examined are:

- Government-Owned, Government-Operated (GOGO)
- Government-Owned, Contractor-Operated (GOCO)
- Government-Owned, Grantee Institution-Operated (GOGIO)

Each is discussed in turn with examples used in other contexts that could be considered for an NSDS.

Government-Owned, Government-Operated (GOGO)

Government-owned, government-operated (GOGO) facilities are owned or leased by the U.S. government and also managed and staffed by employees of the federal government. These GOGO facilities are bound by federal laws, regulations, and policies directed to the federal government and its agencies. They conduct their business representing the federal government. Examples of GOGOs are the U.S. Department of Energy's National Energy Technology Laboratory (NETL) and the U.S. Department of Transportation's (DOT) Volpe Center.

The National Energy Technology Laboratory develops and delivers integrated sustainable energy solutions. NETL is organized into four Centers and one unit under the Office of the Director: The Research and Innovation Center, the Technology Development Center, the Laboratory Operations Center, the Finance & Acquisitions Center, and the Science & Technology Strategic Plans & Programs unit. The centers and unit are located in Albany, New York, Morgantown, West Virginia, and Pittsburgh, Pennsylvania.

The Volpe Center provides multidisciplinary, multimodal transportation expertise on behalf of U.S. DOT's operating administrations as well as other federal agencies, state and local governments, academia, and industry.³¹ Its mission is to advance transportation innovation for the public good by addressing challenges, collaborating, and developing solutions that advance the national and global transportation systems. It is located in Cambridge, Massachusetts.

Government-Owned, Contractor-Operated (GOCO)

Government-owned, contractor-operated (GOCO) facilities are owned or leased by the U.S. government but managed by third-party contractors. GOCO contractors operate the GOCO facility and are subject to the Federal Acquisition Regulations (FAR). The GOCO is run by a contractor that ultimately delivers to the government the products and services in its statement of work.

One example of a GOCO activity is the U.S. Army munitions manufacturing plants. The Army owns several facilities but contracts out the management and production activities. A typical contract may run seven years with options to extend. Government employees are often working in the facility along with the contractors, and these facilities are located throughout the country. At some of these facilities, the contractor is also charged with developing partnerships, community outreach, and other activities. The production plants generally have oversight from the government agency but do not have special governance structures beyond the normal management of the corporation that is awarded the contract.

Another example of a GOCO operation is a Federally Funded Research and Development Center (FFRDC), designed to meet a special long-term research and development need that could not be met as effectively by existing in-house or contractor resources. Several federal laboratories, including 16 of the 17 U.S. Department of Energy National Labs, are FFRDCs.³² The NASA Jet Propulsion Lab is an FFRDC managed by a contractor for NASA.³³ A federal laboratory is “any laboratory, any federally-funded research and development center (FFRDC), or any center that is owned, leased, or otherwise used by a federal agency and funded by the federal government, whether operated by the government or by a contractor.”³⁴

As laid out in FAR 35.017, an agency can use an FFRDC contractor to accomplish tasks that are integral to its mission and operation but could benefit from expanded resources available through the private sector. FFRDCs are designed with agile authorities that enable them to provide agencies with new and cutting-edge scientific support and technical expertise. In total, the government has 42 FFRDCs sponsored by 18 different agencies.³⁵ Five of the FFRDCs are operated by industrial firms, 22 by non-university non-profits, and 15 by universities. Many of the operators are consortiums.

The FFRDCs vary considerably and many developed over time from earlier research efforts that were consolidated into FFRDCs. The FFRDCs generally have mature governance structures that go beyond what one might expect through a regular contract arrangement.

Government-Owned, Grantee Institution-Operated (NOGIO)

The federal government has many different models for awarding grants. There are statutes, OMB policy guidance, and agency requirements related to applying for and receiving grants, and many agencies have specific or directed grant authority. However, only one specific grant program is included in this type of model—the Mathematical Sciences Institutes, established and funded by NSF.³⁶ The five year grants provided by NSF help build math capacity by advancing research, increasing the impact of the mathematical sciences in other disciplines, and expanding the research talent base.

The grants are awarded to institutions every five years, and NSF has the option to renew the grants every five years or move the grant to a different institution. The size of the grant varies among the institutes and they also are allowed to develop additional sources of funding beyond what is provided by the NSF grant. These grants are included in this report’s scope primarily because of their governance structures—the institutes have mature governance structures that comply with a complex set of federal requirements as well as having demonstrated value to NSF and the scientific community.

Application of Potential Governance Structures to the NSDS at NSF

Governance and oversight are key to establishing transparency and trust. The NSDS needs to demonstrate to key stakeholders and oversight bodies that it is operating according to broadly accepted principles and practices that are ethical and equitable, and that the work being produced is independent and of high quality. For the NSDS located at NSF, oversight could be expected from NSF leadership, the National Science Board, the NSF Inspector General, OMB, Congress, and the Government Accountability Office (GAO). But governance needs to come from the stakeholders, including representatives of data providers, owners, and users. The following are federal stakeholders who should have a voice in the overall governance of an NSDS:

- The Interagency Council on Statistical Policy (ICSP/ Chief Statistician of the U.S.)
- Chief Data Officers Council (Chair)
- Evaluation Officers Council (Chair)
- Federal Privacy Council (Chair)
- Chief Information Officers Council (Chair)
- Director of NCSES

Ideally stakeholders outside the federal government would be included to provide input for the NSDS. These should include, at a minimum:

- Senior executives or lead partners from the contractor organization(s)
- State and tribal government representatives
- Privacy expert
- Cybersecurity expert
- Research and evaluation community representative
- Ethics expert

There are several models for how this governance body could be chartered and established. These models assume existing administrative and legal authority provided under current law or through enactment of legislation, such as the NSDS Act, to envision an implemented state.

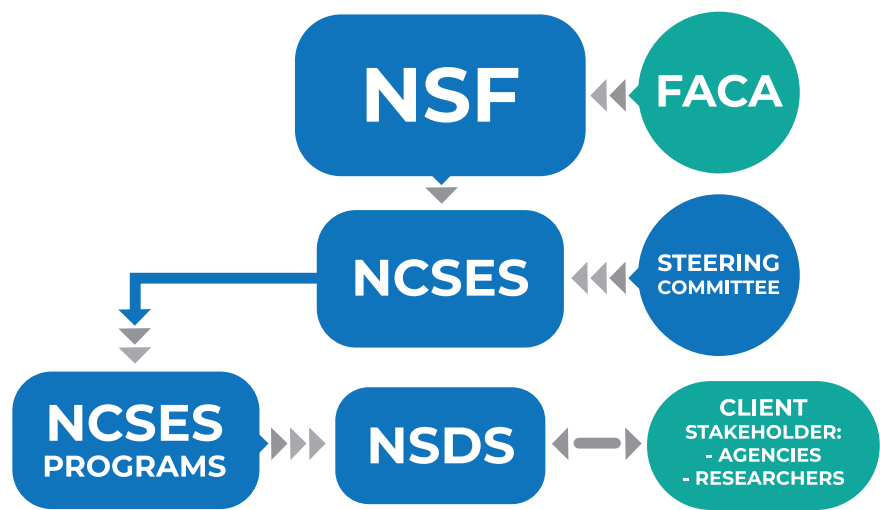
Government-Owned, Government-Operated (GOGO): The GOGO model at NSF could operate within NSF’s current agency organization, but it would be limited by the complicated structures and approaches used to implement NSF programs more broadly within a government-grant making institution. Additionally, the GOGO model has major limitations in its ability to conduct meaningful engagement and outreach necessary to satisfy the attributes discussed previously. For example, routine meetings with stakeholders may trigger requirements under the Federal Advisory Committee Act (FACA) or solicitations for data and information about processes could trigger requirements under the Paperwork Reduction Act. In either case, the approaches could result in more, not less, work to generate relatively basic insights and perspectives that could be gathered in other ways at lower cost. Nonetheless, applying the model at NSF could be reasonably accomplished in one of two ways:

Example 1 – FACA and Steering Committees (Figure 1). Establish a steering committee with the federal members, chaired by the Chief Statistician of the U.S. and engage outside stakeholders through establishment of a Federal Advisory Committee. The bifurcated structure would be necessary because without an explicit statutory provision, the non-federal participants would only be able to serve in an advisory role for governance matters and that would be guided under FACA, whereas federal employees could be members of a steering committee chartered to make decisions. The benefit of this approach is that it could occur relatively quickly at low cost. However, the passive approach to engagement and limited involvement of stakeholders would likely be a challenge. The model is also limited in its ability to identify problems then to rapidly harness the public-private partnership for feedback and iteration to resolve or address potential issues expeditiously as they arise.

Government-Owned, Government-Operated

Example 1

*The Steering Committee is assumed in the Example 1 diagram to be composed of the federal government participants. The Federal Advisory Committee Act (FACA) committee would be non-federal advisors.

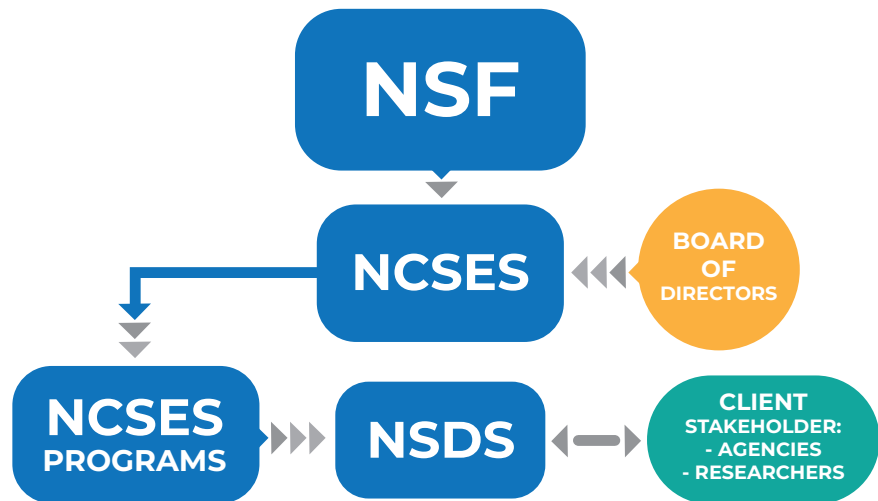


Example 2 – Presidentially-Appointed Board (Figure 2). Establish an NSDS Board similar to the NSF National Science Board (NSB), which includes members and a director.³⁷ The NSB establishes the policies of NSF, identifies issues that are critical to NSF’s future, approves NSF’s strategic budget direction and the annual budget submission to OMB, and approves new major programs and awards. It is made up of 25 members appointed by the President. The NSF Director is an ex officio member. Members serve six-year terms and one-third of the Board is appointed every two years. Board members are drawn from industry and universities and represent a variety of science and engineering disciplines and geographic areas. The board elects its own Chair and Vice Chair, and the Chair is authorized to make appointments to the board staff, headed by an Executive Officer. The NSB meets about five times a year, and the public is invited to attend all open sessions.

**Government-Owned,
Government-Operated**

Example 2

*The Board of Directors in Example 2 is assumed to be composed of both non-federal and federal members who would be able to jointly participate either through a suitable existing legal authority or new authority provided by statute.



If such a board were established, it should include federal members and specify relevant areas of expertise for the presidential appointees. While the full NSF Board is 25 people, a manageable size for an NSDS board would probably be around 10-15 people. This option is likely only possible with separate authorizing legislation that explicitly establishes this structure, unless the current authorization of the NSB would allow it to establish a committee that would serve as the Board of Directors for the NSDS. We did not delve into the limits of the National Science Board’s authorization or its desire to take on responsibility for a significant new committee.

Government-Owned, Contractor-Operated (GOCO): This model has benefits that arise from providing government oversight, direction, prioritization, and accountability on key topics and issues that align with emerging needs and resource availability, while also leveraging the private or non-profit sectors’ ability to hire talent, establish diverse oversight, build partnerships, incorporate innovations, and conduct analyses relatively quickly.

Example 3 – Contractor-Selected Board (Figure 3). Establish a Board of Directors (or Governors) under the aegis of the contractor. Depending on whether the contractor is a single entity or a consortium, the exact makeup of the board would vary. Two or more entities form an association or a limited liability company (LLC) that manages the NSDS under contract to the NSDS. The LLC forms a Board of Directors which enters into the contract on behalf of the LLC. The Board of Directors consists of the key leaders from the LLC companies and outside stakeholders and experts, and the federal agency is usually represented as ex-officio members of the board. ³⁸

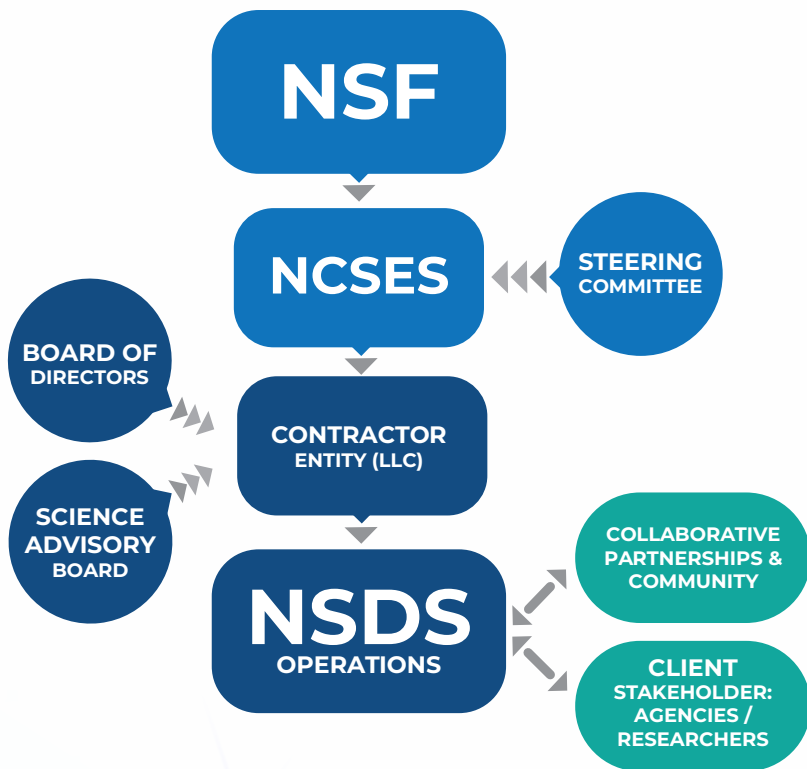
Under this governance structure, the agency determines the mission and provides funding for the entity, and the entity management team conducts the work. The agency generally has an on-site presence. The Board of Directors provides oversight over the entity in areas such as guiding principles, an ethical framework, Diversity, Equity, and Inclusion (DEI), cooperative research and development priority areas, business management and operations, cyber security, outreach and partnerships, infrastructure investment, and strategic planning. Some laboratories have established separate Science Advisory Boards and others have science advisory committees of the Board of Directors. This type of model is currently in use at the Department of Energy’s Lawrence Livermore National Laboratory, Brookhaven National Laboratory, and Los Alamos National Laboratory.

Example 4 – Sole Contractor’s Existing Board (Figure 3). The contractor is a single entity, such as a university, and the designated Board of Directors or Governors then reports to the university president. However, the functions of the board are similar to Example 3. There are some variations on this, for example the board may be a subcommittee of the university’s own Board of Trustees. Generally, though, there is a scientific advisory function present either as a separate committee or as a subset of the Board of Directors. This type of governance model is currently in use at the Department of Energy’s Ames and Argonne National Laboratories, and NASA’s Jet Propulsion Laboratory.

Figure 3. Government-Owned, Contractor-Operated

Example 3 & 4

*This diagram shows how a typical GOCO structure might work with an NSDS. The Board of Directors and Science Advisory Board are run by the contractor entity and do not have federal employees. Rather, the federal employees are on the Steering Committee that provides input to NCSES, which then conveys policies to the contractor.



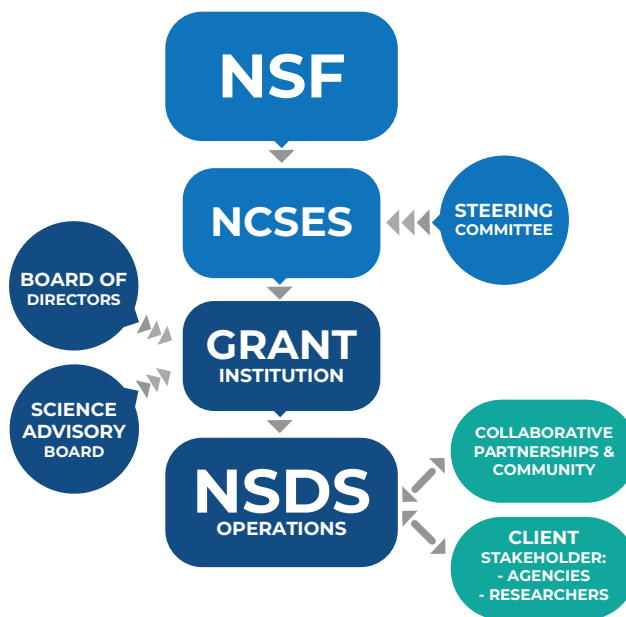
Government-Owned, Grant Institution-Operated (GOGIO)

Example 5 – Grant Institution (Figure 4). The five NSF Mathematical Sciences Institutes have similar, if not identical, governance structures. They each have Boards of Trustees to oversee their activities, Science Advisory Boards that select specific scientific projects to undertake, and Executive Directors.

Figure 4. Government-Owned, Grant-Operated

Institution Example 5.*

*Although the diagram shows the NSDS grant going through NCES, the grants for the math institutes come from a different division of NSF. This portrayal is for illustrative purposes of how a grant might work for the NSDS. Similar to that arrangement in a Government-Owned Contractor-Operated organization, the Board of Directors and Science Advisory Board are run by the Grant Institution and the Steering Committee would be composed of the federal employees advising NCES.



Boards of Trustees. The size and composition of the Trustee Boards vary, but the institutes that are located within a university generally have 8–16 members. Board members, as encouraged by NSF, serve 2 to 4-year terms and are generally limited to two terms, unless serving as the Chair of the Board. The Trustees are elected by the Boards. The Boards of Trustees oversee all institute activities, including reviewing the budget, advising on policies, advising on the appointment of new institute directors, and actively participating in fundraising. Some of the Boards are also engaged in promoting DEI in institute activities. Trustees come from diverse backgrounds in the government, non-profit, academic, and corporate sectors. Most of the Boards of Trustees meet once a year, although working groups may meet throughout the year.

Science Advisory Boards. The size and composition of Science Advisory Boards also varies between institutions, but they generally have between 8–12 members. Board members serve two terms that range from 3 to 4 years. The Science Advisory Board members are elected by the boards. They are responsible for selecting the projects that the institutes will undertake. Board members are highly qualified scientists and mathematicians, predominantly from academia, who can assess the rigor of submitted project proposals. The Boards meet at least annually to review submissions but may meet more frequently and may have interactions with submitters to improve proposals before they are accepted. The programs of the math institutes are varied, and proposals may be to organize in-depth workshops, summer programs geared toward attracting students to work on industry and government agency-sponsored projects, long programs that can last a semester, host special events, and other relevant scientific and educational activities.

Executive Directors. The director of the institute is usually hired as a tenured professor at the grant institute if a university is hosting the institute. Therefore, they are an employee of the university, and the grant goes to the university. All activities at the institutes then conform to the rules and regulations of the university and the policies governing an NSF grant.

The institutes are often governed between meetings by an Executive Committee consisting of the institute Director and the Chairs of the two boards. While the membership is separate, there is communication between the two boards to assure that activities are consistent and complementary. NSF program managers monitor the grants and may attend the board meetings as ex-officio members. However, they are not involved in the day-to-day oversight or decision-making activities of the institutes.

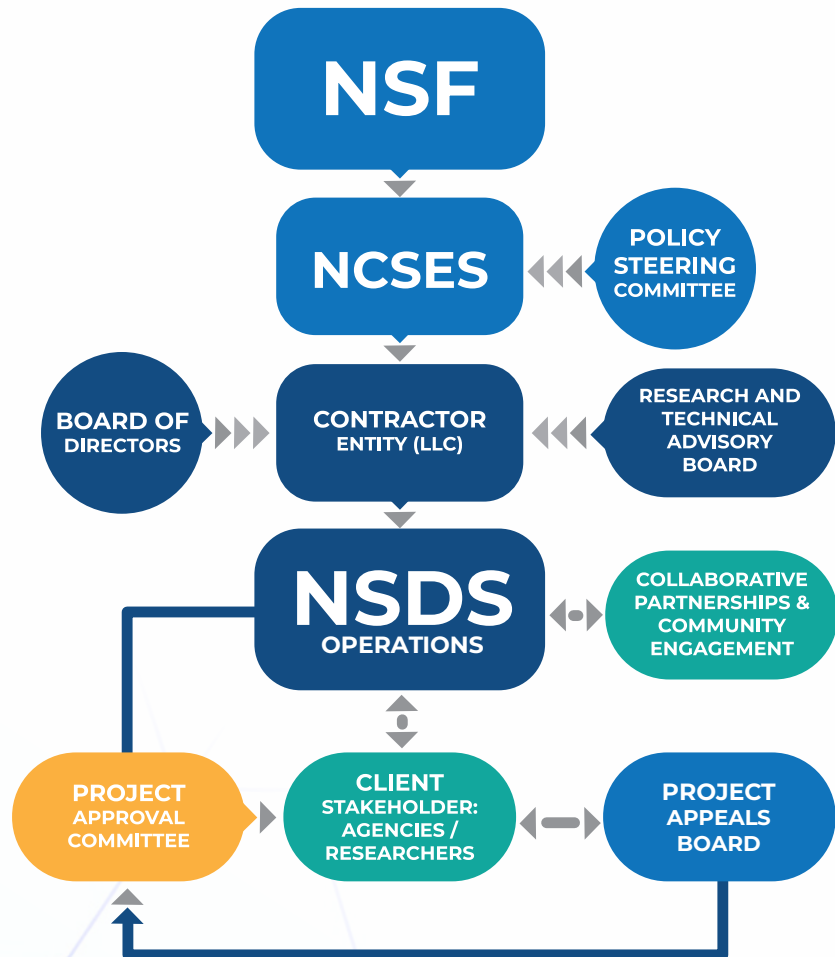
Blueprint for NSDS Governance Model

Each of the above examples have some disadvantages and advantages for an NSDS governance structure. This report proposes a combined structure incorporating attributes from each of the above examples that would best support the features of the NSDS under the government-owned, contractor-operated infrastructure model as proposed in the Data Foundation's *Modernizing U.S. Data Infrastructure* report. Figure 5 is a representation of the various relationships that would be beneficial when utilizing the GOCO blueprint structure.

The summary of the roles and responsibilities of each of the organizational units is included below. This model assumes that the application process for requesting access to data from statistical agencies would be integrated into the application process for the NSDS. It is possible that the ICSP would want to maintain a separate application process strictly for statistical agencies, but in the long-term this may be duplicative, inefficient, more expensive, and could create delays and additional barriers for researchers. As discussed below, the single application process should provide value rather than create delays, and it needs to be adequately resourced to function in a timely way.

Figure 5.
**Government-Owned,
Government-
Contracted**

Recommended blueprint hybrid
governance structure for a GOCO
NSDS



NSF: NSF would provide executive direction to NCSSES, as it does now. NSF would be responsible for all the same leadership and management functions it currently carries out regarding the NCSSES budget, performance evaluation, coordination of cross-cutting activities across NSF, and coordination with the National Science Board. The Chief Data Officer/CIO of NSF would sit on the Steering Committee. NSF procurement would be responsible for managing the administrative aspects of the contract. This implementation approach would cause the least disruption in NSF but still allow NCSSES to be the focal point for the NSDS, a key element of maintaining the protections from the statistical system that are needed to maintain trust, protect sensitive data, and conduct statistical activities that conform to OMB standards and practices. However, NSF leadership, including the NSB, will need commit to the success of this venture as an ongoing service that NSF is stewarding for a larger community, rather than a lower priority contract run by one of its units. Without recognition from NSF leadership that this is a critical, government-wide function and a cornerstone of implementing evidence-based public policy, the NSDS will be significantly hampered in its ability to reach its potential.

NCSSES: NCSSES would be the project owner and manager for the contract. Although the NSDS program would ultimately be the responsibility of the NCSSES Director, NCSSES would need to set up a sufficiently resourced program management office to work closely with the contractor, providing policy guidance, attending board meetings as an ex-officio member, monitoring performance, supporting the Policy Steering Committee, and engaging in other related functions. This structure enables a single point of accountability, allows NCSSES to designate the contractor as a CIPSEA agent (a prerequisite to successful operation), maintains a critical link to the Federal Statistical System, and allows for additional coordination across NSF as needed. However, NCSSES would need to be given a greater voice within NSF for determining the resources needed to manage the program than it currently exercises with its own budget as part of the larger NSF appropriation. That is, the national service aspect of the NSDS should not be subsumed in changing NSF priorities, because the government-wide responsibility should be reliable and agencies will be depending on the longer-term viability of the enterprise.

Policy Steering Committee: The Policy Steering Committee would consist of the federal Executive Branch stakeholders in the NSDS. It would be chaired by the Director of NCSSES. Members would include the chairs of the ICSP, CDO Council, Federal Privacy Council, Evaluation Officer Council, and CIO Council. There could be two to three seats that rotated among the statistical agencies and two to three seats that rotated among program agencies which may be supplying data to the NSDS. The Policy Steering Committee would be responsible for advising the NCSSES Director on significant policy issues around research priorities, protecting sensitive data, cybersecurity, data quality, and other relevant policy concerns. It would be the responsibility of the NCSSES Director to appropriately convey the policy direction to the contractor. This structure enables key input from the data owners as well as high level officials with a cross-agency perspective, such as the Chief Statistician of the U.S., the Federal CIO, the Lead Privacy Official in OMB, and others.

Contractor Entity: The contractor entity would consist of the organization or multiple organizations that successfully competed for the NSDS contract. This could be a consortium of universities, a combination of non-university and university non-profits, a for-profit corporation, with or without partners, or various combinations. The entity, most likely an LLC, would be responsible for, but not limited to, the following activities, including: (1) assuring that the research and operations met the contract requirements, complied with policy directives from NCSSES, and with federal law and regulations; (2) hiring the Executive Director of the NSDS; (3) establishing and interacting with both the Board of Directors and the Research and Technical Advisory Board; and (4) being the primary point of contact for NCSSES leadership. This structure assures that there is a single point of accountability on the contractor end for accomplishing the scope of the contract and maintaining a schedule of deliverables that are documented consistently and in a way they are able to be audited. This approach puts one entity (that could consist of multiple partnering organizations) in charge of the contract but does not preclude participation for a national network of nodes operating independently or through subcontract awards, or a combination of these. It allows the flexibility needed to be responsive to the public and community of users but still maintains overall government control. The entity needs to consist of an organization or organizations that have a deep understanding of social science research and data science.

Board of Directors: The Board of Directors would consist of the senior executives of the contractor entity and members of the general public. The Board should include privacy, cybersecurity, and ethics experts, community leaders, researchers, and experienced executives and managers. Members would initially be selected by the entity, but the board would eventually select its own members under the guidance of the Chair. It would be responsible for providing oversight in areas such as guiding principles, an ethical framework, DEI, business management and operations, cybersecurity, outreach and partnerships, infrastructure investment, and strategic planning. The board would consist of between 8–24 members serving 3–4-year terms that could be renewed once. The board would meet at least once annually and be chartered to set up working groups as it deems necessary. This structure is a model that has worked for multiple organizations that need to be accountable to the public.

Research and Technology Advisory Board: The Research and Technology Advisory Board would provide advice to the entity on research directions, technology advances that should be explored, statistical methodology around data linkage and data quality, and other related areas. The members of the board should be top experts in these fields from academia, non-profits, and the private sector. The board would consist of 8–12 members with complementary expertise and serve 3 to 4 year terms that could be renewed once. Members would initially be selected by the entity with staggered terms but would eventually select its own members under the guidance of the Chair. The board would meet at least once annually and be chartered to set up working groups as it deems necessary. This type of board is a model that has worked well for many scientifically-oriented public enterprises.

NSDS Operations: The NSDS operations functions would be led by an Executive Director selected by the Contractor Entity with advice from the Board of Directors. The operations would include all the management and operations of the functions taken on by the NSDS. The contractor would need to hire appropriate staff, manage the budget, work with all stakeholders on an on-going basis, including agencies providing data, researchers requesting data, and collaborative partners. Quality control, privacy protection, cybersecurity, internal research, infrastructure procurement and maintenance, communications, and other like functions would be under the management of the Executive Director. This structure is key to successful implementation of service provision. It allows operational decisions to be made and efficiently executed under the broader policies and principles set by the governing bodies.

Collaborative Partnerships and Community Engagement: To be a transparent and trusted organization, the NSDS would need to undertake broad partnerships with state, tribal, and local governments, other research institutions, and with communities that coalesce around social science evidence-building topics that affect public policies. While these collaborations and engagements will vary, they are specifically included because they are an important feedback mechanism for the NSDS as well as have a multiplier effect in strengthening capacity for evidence-building. There will need to be a part of the NSDS that develops these partnerships and collaborations. The value of having sustained focus on developing federal-state partnerships around data cannot be overstated.

Client Stakeholders – Agencies and Researchers: Agencies providing data and researchers accessing data are the primary clients of the NSDS operations. They will need workstreams that are efficient, reliable, and timely to access data for evidence-building. The NSDS will need to pay close attention to its business processes for carrying out these functions, starting with the application process and going through to data dissemination. This aspect of the overall structure needs to include a feedback mechanism from users and responsive client relationship management.

Project Approval Committee: This body is one of the most critical functions of an NSDS. It would consist of experts appointed by their agencies, who are delegated the authority to meet and approve or disapprove projects according to a set of criteria that have been thoroughly vetted to be in conformance with guidance from the Policy Steering Committee and NCSES. The criteria should be publicly available for transparency. One question to be addressed is whether a committee should approve all projects, such as if a researcher wants data only from one agency or if the committee should approve only projects requesting linked data from multiple agencies. The Project Approval Committee could set up expedited processes for less complicated data requests. For example, a fast-track project approval may consist of a request for data from one agency and may need review by only one member of the committee from the agency from which data are being requested, rather than waiting for the committee to convene.

There would not have to be a one-size-fits-all business process for approvals. However, it would be essential for the committee to meet as a group at least biannually – preferably quarterly – to review proposals that require data from multiple agencies. If project approval became a bottleneck, the NSDS and the evidence-building process as a whole would not successfully achieve the goals of the Evidence Act. Part of the responsibilities of the committee would be to assure dedication and focus on the part of all the committee members. The committee would have a charter and members could have three-year renewable terms, subject to reappointment by their agency. The project approval committee would be chaired by a member of the committee named by the Director of NCSES, in consultation with the Policy Steering Committee. The committee’s work would be supported by the NSDS operations contractor, who would be charged in their scope of work with providing sufficient resources to maintain an efficient, reliable, and timely business process for approval consideration. This would help the smaller agencies that often cannot muster sufficient resources to approve requests for access to data in a timely manner.

Discussed in more detail in the business processes section below, this report acknowledges the significant work that has been done by the ICSP in looking at a project approval structure for the application portal mandated by the Evidence Act. It is hard to imagine, though, a separate application and approval process for projects requesting data only from statistical agencies and projects that use linked administrative data from non-statistical agencies. Having parallel application and approval processes would create additional barriers for applicants and has the potential to create confusion and unnecessary costs and redundancy. The application processes should be combined at the NSDS to the extent practicable.

Project Appeals Board: The Project Appeals Board has a statutory function to consider projects that have been turned down if the requestor would like to appeal that negative decision. A three-person standing board that meets quarterly to consider appeals would comply with this mandate. There would be two members appointed to three-year terms and a third member who would rotate on the board to consider appeals for data from their agency. That is, two permanent members and one member who changes depending on the nature of the appeal. If the project being appealed involves data from multiple agencies, one representative would be selected. The members of the appeals board with fixed terms would be appointed by the Director of NCSES in consultation with the Policy Steering Committee.

Business Processes for Operating the NSDS

Efficient, transparent, and timely business processes are essential elements for a successful NSDS. One of the primary reasons the 2020 *Modernizing U.S. Data Infrastructure* report recommended that the NSDS be a Government-Owned Contractor-Operated facility was to enable it to dedicate sufficient resources and hire and retain highly qualified employees more easily than seen with typical federal government processes. In addition, a contractor can exercise flexible authority in working with partners, building collaborative communities and cooperative arrangements around particular topics of interest, and bring in research fellows and students to build pipelines of future federal employees in the data and statistical agencies.

The business processes for an NSDS would provide authorized access to sensitive data, potentially link data from multiple sources, protect the data throughout its stay in the NSDS, and assure that identities would not be disclosed when the research results are disseminated.

Mentioned above, as required by the Evidence Act, a single application and process for requesting access to data from a statistical agency is being developed.³⁹ Of course, as noted, much data used for evidence-building activities are program data collected by non-statistical agencies that are not mandated by law to use a single application or a standard government-wide process to consider proposals and grant access to data to researchers. Many of these agencies have long-standing processes and ongoing relationships with the research community. To make an NSDS attractive as a government-wide service, the business processes have to offer value beyond what non-statistical agencies can accomplish on their own. Additionally, most of the statistical agencies currently have arrangements in place to share data with each other under different authorities and use a variety of different facilities.

There are multiple facilities, run by multiple organizations, where access is granted to confidential data. These facilities include Federal Statistical Research Data Centers, several centers located at universities, and a handful of non-profit organizations. A goal of the Evidence Act is to continue to build capacity, not create a monopoly for data sharing and linkage. Although the Evidence Act requires a single application and approval process, it does not require that actual access to analyze data take place at a single facility—which is not practical at scale.

Some existing facilities specialize in specific policy topics, such as health or poverty, training, and have existing relationships with states or with universities that may be the home institution of the researcher. The NSDS is an ideal place to host the application and approval process and to help facilitate locating the research at the facility that would provide the optimum benefit to agencies and researchers. The NSDS should have the capacity to provide researcher access and conduct data linkages and disclosure avoidance activities, but it could also serve as a coordinating body to encourage capacity-building across multiple institutions and facilities.

A major question that remains for defining business processes and purpose is: What would be considered valuable activities for the NSDS? Below are a few of the attributes that are already being developed or might be worthwhile to develop further:

1. An easy-to-understand and easy-to-use application for requesting data (currently being developed for statistical agencies)⁴⁰
2. A fast, transparent approval process, even if data were requested from multiple agencies, including any needed background checks
3. Easy access to prerequisites for using data, such as agency-required data stewardship training, licensing agreements, and nondisclosure oaths.
4. Infrastructure to conduct data analyses, including secure remote access
5. A fast, transparent process for undergoing disclosure avoidance reviews before release of research results
6. An easy-to-understand equitable cost structure for using the NSDS services
7. Online access to information about the datasets, such as what other research has been conducted using these data; how to connect with researchers who have used the data; information about the data quality; related data sets; and standard metadata elements
8. Standard data sharing agreements that are “off-the-shelf” for agencies to share data among themselves
9. A library of best practices surrounding various aspects of using specific data sets
10. Information about and access to a network of researchers, experts and facilities across the nation that are pre-approved to host NSDS-approved projects, such as academic big data centers, non-profits, social science research centers affiliated with academia, state collaboratives, and other like organizations. The goal would be to utilize and build capacity across multiple institutions and use the NSDS to keep track of and coordinate communities of interest
11. Opportunities to participate in exciting research and educational activities with other researchers, social scientists, and data scientists, such as workshops, fellowships, information exchanges, cutting edge research projects, and technical training for working with data at all stages of one’s career
12. The capacity to conduct these activities could be added over time, but the essential core business processes would need to be in place for some of the added-value activities to be well received. It will be important for the NSDS management to identify where agency policies are creating bottlenecks and assist agencies in finding more efficient approaches to meeting their requirements.

Even with these many considerations for business processes, there are still yet other questions to consider. The NSDS will not be able to undertake all activities envisioned and—like all organizations—must determine how to partner, collaborate, or delegate. There will likely also be technical questions that arise about hosting of projects, privacy technologies, curation, and the many other critical components for operational success. Our advice is that the NSDS rely on the robust advisory processes established in our recommended blueprint, and building on the advice of ACDEB, to make rapid advances on these important issues, while aligning with the intent of the Evidence Act.

Sustainable Funding Models

Sustainable funding is imperative for the success of the NSDS. There are multiple ways in which the NSDS could be funded. However, at the core is a federal appropriation to NSDS, which would then be used to fund a contract with a contractor entity and to pay for program management at NCSES. The NSDS demonstration project legislation that passed the House authorized \$45 million over 5 years (\$9 million per year) to operate directly or contract out the NSDS.⁴¹ The FY 2023 President's Budget request also includes references to the need for direct funding for the data service activities.

Another aspect of funding is user charges, which OMB has existing guidance and expectations about how to collect in many cases for certain services.⁴² There are multiple models for charging for these types of services, including some that the Evidence Commission considered in its design for the NSDS.⁴³ One easy-to-understand model could be based on a flat fee for the approval process, fees for computation time that include costs of disclosure review, data preparation, etc. There could also be a menu of optional services with associated fees as well. The fees would not be prohibitive and would be structured so that if a Principal Investigator had multiple graduate students accessing data for a research project, the costs would not be unreasonable. This approach would also be consistent with the recommendations and comments from the Evidence Commission about implementing the NSDS.⁴⁴ Cloud-based infrastructure would keep capital investment costs low and enable easy tracking of usage.

A third source of revenue for the NSDS would be through special projects and partnerships. The NSDS contractor would be able to work under cooperative agreements and have sponsored events paid for by partners. For example, the Frederick National Laboratory works with three major agreement types: (1) Material Transfer Agreements (MTAs) that facilitate the exchange of materials for research or testing between the lab and a partner; (2) Collaboration Agreements that enable intellectual and material contributions and unfunded collaborative work between the laboratory and one or more partners; and (3) Contractor Cooperative Research and Development Agreements (cCRADAs) that enable the laboratory to collaborate with one or more partners on jointly conducted research and accept cost-recovery funding from partners. Intellectual property and other research outcomes are shared among the partners.

Timing for Establishing an NSDS

If this blueprint is followed, an NSDS could begin operating in 12–24 months from launch time. Many of the activities for establishing the NSDS could run in parallel. Realistically, time from appropriation of funds to award of the contract could take 6-12 months. This compares to the time to organize internally to establish an operational NSDS as a GOGO facility, including hiring staff, finding or building the infrastructure, setting up the governance, etc., which could easily take much longer than 12 months.

The NCSES should begin to structure the competitive bidding for an NSDS pilot now so that an RFP could be released shortly after enactment of authorizing legislation and NCSES receipt of appropriations. At the same time that the RFP is out for bid, NSF and NCSES should begin organizing the governmental structures for governance. NCSES should examine closely whether America's Data Hub could be used to initiate parts of this process, particularly as it relates to beginning to coordinate research nodes, which would also save considerable time in establishing NSDS as a coordination point for a national network of research facilities involving federal data linkage. First, though, NCSES needs to establish a properly resourced project management office. Then, working with OMB, the Policy Steering Committee

could be established, and the ICSP could structure the NSDS application approval process around the substantial work that has already been done to implement the statistical agency pilot application portal. This work would need some minor tweaking to be adapted to non-statistical agencies. In addition, the Policy Steering Committee could create the first drafts of charters for the Board of Directors and the Research and Technology Board.

Once the contract is awarded, the contractor would have the responsibility for setting up the infrastructure and bringing on staff. The timelines and specifics should be included in the contract. Once the contract is awarded, the contractor entity can also fill the positions on the Board of Directors and Research and Technology Boards and finalize their charters. This would be done in consultation with NCSES.

By carrying out many of the planning and governance implementation functions while the contract is being competed, utilizing America’s Data Hub where appropriate, the time to implementation can be shortened considerably. With input from the ACDEB, the time to begin is now.

Table 1. Summary comparison of different models

Governance Functions	GOGO	GOCO	Grantee	Blueprint
Board of Trustees/ Directors	No	Yes	Yes	Yes
Science Advisory Board	Yes	Yes	Yes	No
Executive Direction	Federal Executive	Executive Director hired by contractor entity	Tenured Professor Director	Executive Director hired by contractor entity
Policy Steering Committee	Yes	Yes	Yes	Yes
Public Participation	FACA	Boards	Boards	Boards and Direct
Research and Technology Advisory Board	No	No	No	Yes
Project Approval Committee	Unknown	No	No	Yes
Operations	GOGO	GOCO	Grantee	Recommended
Daily Operations and management	Federal employees	Contractor entity employees	Grant institution employees	Contractor entity employees
Hiring Authority	Civil Service	Consistent with contractor procedures	Consistent with university procedures	Consistent with contractor procedures
Partnership Development	Federal agency led	Contractor entity led	Grantee institution led	Contractor entity led
Business Process development and execution	Federal agency led	Contractor entity led	Grantee Institution led	Contractor entity led with input from NCSES and policy committee
Sustainable Funding	GOGO	GOCO	Grantee	Recommended
Federal Appropriation	Yes	Yes	Yes	Yes
Cost recovery through fees	Unknown	Unknown	Unknown	Yes
Partnerships/Sponsors	Unknown	Yes	Yes	Yes

Summary Findings

Although first launched as a pilot project, the initial structure and functions of the NSDS should put a foundation in place that could both continue to serve the needs of a core NSDS and be flexible enough to add additional functions if they are viewed as valuable as the NSDS matures.

- There is no recommended timeframe for when the NSDS can be considered mature—if successful, it should demonstrate value and continuous learning that gains a consensus among its users that adding functionality is beneficial.
- The NSDS should adopt a hybrid governance structure incorporating the attributes that would best support the functions of the NSDS under the infrastructure model proposed in the first paper, that is, a government-owned, contractor-operated facility.
- Under a GOCO model, there would be specific roles for the various aspects of the governance structure, including for NSF, NCSSES, a Policy Steering Committee, the Contractor Entity, a Board of Governors, a Research and Technology Advisory Board, NSDS Operations, Collaborative Partnerships, Client Stakeholders, a Project Approval Committee, and a Project Appeals Board.
- Efficient, transparent, and timely business processes are essential elements for a successful NSDS. The NSDS should not become a bottleneck. The NSDS would be an ideal place to host the application and approval process, provide core services, and help facilitate locating the research at other facilities if that would provide the optimum benefit. However, the NSDS management will need to identify where agency policies are creating bottlenecks and assist agencies in finding more efficient approaches to meeting their requirements.
- Sustainable funding is imperative for the success of an NSDS. The base funding should come from a federal appropriation to NCSSES, which would then be used to fund a contract with a contractor entity and to pay for program management at NCSSES. Funding should also come from user charges, which should be easily understandable and published on an equitable fee-for-service schedule based on usage that does not constitute a barrier to access for researchers. The contractor should be able to also collect revenue as part of partnerships and cooperative agreements.

Recommendations for Next Steps

- The NCSES should begin to structure the competitive bidding for an NSDS pilot so that a Request for Proposals could be released shortly after enactment of authorizing legislation and NCSES receipt of appropriations as authorized in FY 2022. NCSES should utilize America's Data Hub where appropriate to move forward portions of the NSDS concept, such as coordination of data linking research centers that currently exist.
- While the competitive bidding process is underway, NCSES should work with OMB to establish the Policy Steering Committee and other governance structures.
- The ICSP should seriously consider structuring the substantial work that has already been done to implement the statistical agency pilot application portal to adapt it to the NSDS application approval process.

There are many questions that will remain left unaddressed about how to conceptualize, establish, launch, staff, operate, interact with, and manage the NSDS at NSF. We did not set out to address every conceivable question in this blueprint, but rather to identify a starting point for NSF on the major questions of governance and operations. We intend for this to be a productive starting point for NSF and to inform the deliberative process of the ACDEB members. Our findings, building from our prior research, suggest a potential path that is reasonable, and a strong improvement from the status quo.

Even as the NSDS becomes a reality through implementation of our proposed blueprint, there will inevitably continue to be challenges for the data community. We expect these challenges will likely arise in navigating the legal frameworks, cooperating across agencies, ensuring the strongest privacy protections are applied, navigating tiered access protocols, and deploying privacy enhancing technologies to the extent feasible, among other areas. These challenges are not insurmountable, but require careful coordination, capacity, leadership, and resources. This is all in the context of operating with a high level of transparency and accountability to assure the American people their data are protected, secure, and ethically used.

There are hundreds, if not thousands, who have already contributed to the design of the NSDS through the Evidence Commission, ACDEB, congressional dialogue, professional societies, and other processes. In the years ahead, there will be thousands more who engage to ensure the NSDS fulfills its active purpose, achieves value, and satisfies its mission. While we offer this blueprint as a starting point, the ultimate success of the National Secure Data Service rests with the data and evidence community.

The statistical system, the evidence community, and the country will all be better off with the NSDS than without it. We will be able to gather better insights across policy domains in new ways. We will be able to answer questions we never have before. And if we do this right, we will be able to safely operate the concept of evidence-based policymaking at a tremendous scale in the U.S. When we succeed with that goal, we will have new tools to use evidence for combating poverty, improving employment outcomes, increasing economic prosperity, and addressing other policy objectives.

Advancing implementation of the National Secure Data Service is a safe, practical, and necessary way to ensure our country's leaders will have the data and evidence they need to make important decisions. The blueprint we offer in this report for NSF proposes a practical strategy for prioritizing how NSF could begin to implement the NSDS in the near-term. Echoing the sentiment of the Evidence Commission – it's time for the vision to become reality.

Resources

1. U.S. Commission on Evidence-Based Policymaking (CEP). The Promise of Evidence-Based Policymaking: Final Report of the Commission on Evidence-Based Policymaking. Washington, D.C.: Government Publishing Office, 2017. Available at: <https://www.datafoundation.org/s/Report-Commission-on-Evidence-Based-Policymaking.pdf>.
2. Foundations for Evidence-Based Policymaking Act of 2018, P.L. 115-435, Jan. 14, 2019. Sec. 101 § 315.
3. Advisory Committee on Data for Evidence Building (ACDEB). Year 1 Report. Washington, D.C., October 2021: p. 1. Available at <https://www.bea.gov/system/files/2021-10/acdeb-year-1-report.pdf>.
4. Advisory Committee on Data for Evidence Building (ACDEB). 2022. Available at www.bea.gov/evidence.
5. N. Hart and N. Potok. Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building. Washington, D.C.: Data Foundation, 2020. Available at: <https://www.datafoundation.org/modernizing-us-data-infrastructure-2020>.
6. Hart and Potok, 2020.
7. CEP 2017.
8. CEP 2017, p. 1.
9. See the compendium of deliberative materials made available by the Evidence Commission as part of its public record available from the Data Foundation at: <https://static1.squarespace.com/static/56534df0e4b0c2babdb6644d/t/5efa3b3cb1e1f2429948381b/1593457475789/Compendium-of-CEP-Staff-Decision-Memos+%281%29.pdf> .
10. P.L. 115-435.
11. P.L. 115-435. Sec. 101 § 315.
12. ACDEB 2021, p. 1.
13. ACDEB 2022.
14. Hart and Potok, 2020.
15. Hart, N. Design Considerations for Implementing a National Secure Data Service. Presented at the Advisory Committee on Data for Evidence Building Public Meeting, April 23, 2021. Available at: <https://www.bea.gov/system/files/2021-04/Hart-NSDS-Framework-for4-23-ACDEB.pdf>.
16. U.S. House. National Secure Data Service Act of 2021 (H.R. 3133). 117th Congress, 2022. Available at: <https://www.congress.gov/bill/117th-congress/house-bill/3133>.
17. U.S. House. U.S. Innovation and Competition Act of 2021 (H.R. 4121). 117th Congress, 2022. <https://www.congress.gov/bill/117th-congress/house-bill/4521>.
18. Abraham, K.A. and Haskins, R. Letter to Congress Endorsing the National Secure Data Service Act. November 30, 2021. Available at: <http://www.datacoalition.org/wp-content/uploads/2021/11/CEP-Co-Chair-Letter-re-NSDS-11-30-2021.pdf>
19. Data Coalition. Fact Sheet on the National Secure Data Service Act. Washington D.C.: Data Foundation, 2021. Available at: <https://www.datacoalition.org/fact-sheet-national-secure-data-service-act-advances-responsible-data-sharing-in-government/>.
20. H.R. 3133, 117th Congress, Sec. 2(a).
21. H.R. 3133, Sec. 2(b)(2).
22. ACDEB 2021.
23. ACDEB 2021, p. 1.
24. America's DataHub Consortium. 2021. Available at: <https://www.americasdatahub.org/>
25. Boyea, K. Presentation to the Advisory Committee on Data for Evidence Building in the January 2022 Public Meeting. Available at: <https://www.bea.gov/system/files/2022-01/Structural-Options-for-America-DataHub.pdf>
26. 42 U.S.C. § 1862(a)(1) and 42 U.S.C. § 1870(c).
27. 27 Young, S. Evidence-Based Policymaking: Learning Agendas and Annual Evaluation Plans. Memorandum for Heads of Executive Departments and Agencies (M-21-27). Washington, D.C.: White House Office of Management and Budget. Available at: <https://www.whitehouse.gov/wp-content/uploads/2021/06/M-21-27.pdf>
28. Note one of this report's authors, Nick Hart, is also a member of the Advisory Committee on Data for Evidence Building.

29. National Academies of Sciences, Engineering, and Medicine (NAEM) Committee on National Statistics (CNSTAT). Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps. Washington, D.C.: The National Academies Press, 2017b. Available at: <https://doi.org/10.17226/24893>; CEP 2017; and ACDEB 2021.
30. Hart and Potok, 2020.
31. US Department of Transportation's Volpe Center.
32. Department of Energy National Laboratories.
33. NASA Jet Propulsion Lab.
34. 15 U.S.C. § 3703.
35. A list of all federal FFRDCs is maintained by NSF at <https://www.nsf.gov/statistics/ffrdclist/#agency>.
36. Note one of the authors of this report, Nancy Potok, chairs the Board of Trustees for a Mathematical Sciences Institute.
37. National Science Board. Washington, D.C., 2022. Available at: <https://www.nsf.gov/nsb/about/index.jsp>.
38. If at some point it was deemed that the federal representatives should officially serve on the LLC board rather than be ex-officio, that may need to be authorized by statute. For example, the composition of the Amtrak Board of Directors, which includes federal employees, Amtrak executives, and outside experts appointed by the President is specified in Title III, Section 303 of the National Rail Passenger Act of 1970 (P.L. 91-518).
39. Interagency Council on Statistical Policy, Office of Management and Budget. Standard Application Policy & ICSP Recommendation for SAP for Requesting Access to Certain Confidential Data Assets. Federal Register. Jan 2022. <https://www.federalregister.gov/documents/2022/01/14/2022-00620/the-interagency-council-on-statistical-policies-recommendation-for-a-standard-application-process-sap>.
40. ICSP 2022.
41. Sec. 8(e)(8) in U.S. House. National Science Foundation for the Future Act (H.R. 2225). 117th Congress, 2022. Available at: <https://www.congress.gov/bill/117th-congress/house-bill/2225>.
42. OMB Circular A-25: User Charges. Available at: <https://www.whitehouse.gov/wp-content/uploads/2017/11/Circular-025.pdf>.
43. See Note 9.
44. CEP, 2017. See also discussion about user charges in the compendium of background files for the commission recommendations.



The Data Foundation is a non-profit think tank based in Washington, D.C., that seeks to improve government and society by using data to inform public policymaking. Our research and educational activities proactively and rapidly address relevant, emerging data-related needs in the country with the goal of devising realistic solutions, accelerating policy coordination, and advancing innovation. The Data Foundation values diversity and transparency in pursuit of an equitable, data-informed society.