



# Advisory Committee on Data for Evidence Building: Year 2 Report

October 14, 2022





October 14, 2022

Shalanda Young, Director of the Office of Management and Budget (OMB)

Dear Director:

On behalf of the Advisory Committee on Data for Evidence Building and pursuant to the Foundations for Evidence-Based Policymaking Act of 2018 (Public Law 115-435), I am pleased to submit the following report detailing the Committee's final findings and recommendations.

This report fulfills the Committee's charge to review, analyze, and make recommendations on how to promote the use of federal data for evidence building. Specifically, the Committee focuses on expanding access to data for evidence building, facilitating data sharing, enabling data linkage, and developing privacy-preserving techniques. The Committee's guidance supports the OMB Director in carrying out duties under Title III of the Evidence Act by providing advice on the current state of evidence building and laying steppingstones for the future.

First, the Committee's recommendations seek to advance implementation of the Evidence Act by informing the forthcoming CIPSEA regulations and providing guidance on other Evidence Act items. In addition, the report casts a vision for the role of the National Secure Data Service—that is, as a new entity entering the evidence ecosystem that provides coordination and capacity-building services. The data service will build on the framework of the Evidence Act and supplement, not replace, the work of other evidence enablers and users. The report maps out the data service's organizational structure, technical infrastructure, and functions to actualize this vision. Finally, the Committee provides a set of recommendations on resources for existing actors in the ecosystem and for the data service itself.

The backdrop of the Committee's work is an existing evidence ecosystem that continues to evolve. Data infrastructures, analytical approaches, privacy-preserving technologies, and other advances in the field are progressing rapidly. Increasingly, data providers and users are demanding services and resources that support their efforts to unleash the power of data and evidence. The legislative and budgetary landscapes also continue to take shape with new initiatives and funding requests that signal support for the federal statistical system's development of core evidence-building infrastructure and capacity.

*Continued*

Even my role on this Committee echoes changes in the surrounding ecosystem. As you know, about 18 months ago, Dominic Mancini, the previous Chair and Acting Chief Statistician of the United States, delegated the Chair position to me, so he could focus on other OMB responsibilities. In May 2022, we welcomed Karin Orvis as the new Chief Statistician. At that time, she asked me to stay on as Chair, rather than take on the role herself. This provided continuity in the Committee's process and maintained a level of independence as the Committee finished its work. I want to thank both Dom and Karin for their invaluable support for the Committee and their broader vision for enhancing evidence-based policymaking.

Over the last 2 years, the Committee amassed a wealth of knowledge by harnessing members' diverse expertise and experience and by engaging with other experts across the public and private sectors. This evidence forms the basis of the Committee's findings and supports the timely, actionable, and relevant recommendations in this report.

The Committee's recommendations provide key guidance at a critical moment, marking the next chapter for evidence building. I call on the Director of OMB to adopt these recommendations and work with stakeholders at all levels of government and throughout the private sector to continue to advance the data evolution.

Respectfully yours,



Emilda Rivers

ACDEB Chair on behalf of fellow Committee members

Laila Alequresh  
Richard Allen  
Otis Brown  
Leonard Burman  
Charles Cutshall  
Shawn Davis  
Gregory Fortelny  
Nicholas Hart  
Christine Heflin

Anna Hui  
Barry Johnson  
Ted Kaouk  
Elisabeth Kovacs  
Edward Kwartler  
Julia Lane  
Christin Lotz  
Brian Moyer  
Kimberly Murnieks

Amy O'Hara  
David Park  
Todd Richardson  
Matthew Soldner  
Kenneth Troske  
Mayank Varia  
Christina Yancey

# Contents

<b>1. Executive Summary</b> .....	<b>1</b>
ACDEB’s Final Report. ....	1
Recommendations Summary .....	2
<b>2. Background</b> .....	<b>6</b>
The Commission on Evidence-Based Policymaking .....	6
The Foundations for Evidence-Based Policymaking Act .....	6
The Evolving Data and Evidence Ecosystem. ....	11
ACDEB’s Purpose, Progress, and Promise .....	15
<b>3. Recommendations</b> .....	<b>18</b>
Introduction: Foundations and Opportunities. ....	18
Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance .....	21
Recommendations Part 2. Other Evidence Act Items .....	39
Recommendations Part 3. NSDS Functions. ....	41
Recommendations Part 4. NSDS Organizational Structure and Governance. ....	59
Recommendations Part 5. NSDS Technical Infrastructure and Tools .....	78
Recommendations Part 6. Resources and Funding .....	90
<b>4. Appendices</b> .....	<b>104</b>
Appendix A: ACDEB Committee Charter, Membership, Process, and Meetings .....	104
Appendix B. ACDEB Use Cases .....	109
Appendix C. ACDEB Virtual Site Visits .....	111
Appendix D. OMB/ICSP Workstreams .....	112
Appendix E. ACDEB Subcommittee Guest Speakers .....	114
<b>References</b> .....	<b>116</b>
<b>Acronyms</b> .....	<b>124</b>

# Executive Summary

## ACDEB's Final Report

The Advisory Committee on Data for Evidence Building (“the Committee” or ACDEB) was established pursuant to the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act). The Committee’s primary task under the law is to review, analyze, and make recommendations to the White House Office of Management and Budget (OMB) Director on how to promote the use of federal data for evidence building, with a particular emphasis on assisting OMB in implementing Title III of the Evidence Act. This task included evaluating the continued need for and value of potential functions and services of the National Secure Data Service (NSDS or “data service”).

The Committee’s first report ([Year 1 report](#)), issued in October 2021, articulated a vision for the NSDS as a philosophy, a service, and a place, and for the future of data sharing, data linkages, and privacy-enhancing techniques across federal agencies and with state, territorial, local, and tribal governments. The Year 1 report provided an initial set of recommended actions to begin building toward that vision and mapped out the Committee’s path for its second year.

Over the past 12 months, the Committee developed its final recommendations by:

- Engaging federal leaders implementing Title III of the Evidence Act;
- Continuing to leverage and share the expertise of its members;
- Hearing from researchers, government leaders, other experts, and the public;
- Conducting additional exploration into real-world projects, use cases, and examples; and
- Collaboratively synthesizing different perspectives and use cases into a coherent understanding of the current state and future needs around the use of data for evidence building.

This report reflects the culmination of the Committee’s 2 years of work to address a very complex set of topics that has been further complicated by a broader, rapidly evolving data and evidence ecosystem. The Committee’s overarching objective for this report is to deliver timely, actionable, and relevant recommendations, fully supported by detailed and nuanced findings that reflect the knowledge and expertise the Committee has shared.

## Recommendations Summary

Rec. #	Recommendation
<b>Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance</b>	
1.1	The Office of Management and Budget (OMB) Director, in coordination with the Interagency Council on Statistical Policy (ICSP), should issue guidance and regulations required under the Evidence Act, incorporating feedback from ACDEB's Year 1 report, subsequent engagements with the Committee, and this report.
1.2	OMB, in coordination with the ICSP and other relevant federal councils, should promulgate required Evidence Act regulations and guidance with an eye toward supporting the NSDS functions, as envisioned by the Committee.
<b><i>Primary Responsibilities of Statistical Agencies and Trust</i></b>	
1.3	OMB should apply practical principles for promoting public trust and should empower all federal agencies to participate in this process.
<b><i>The Presumption of Accessibility for Statistical Agencies and Units</i></b>	
1.4	OMB should issue the Presumption of Accessibility rule to maximize the impact of federal administrative data for evidence-based decisionmaking.
1.5	OMB, in coordination with the ICSP and other relevant federal councils, should identify mechanisms for streamlining data-sharing agreements across federal agencies.
<b><i>Expanding Secure Access to CIPSEA Data Assets</i></b>	
1.6	OMB should adopt a risk-utility framework as the basis for standards on sensitivity levels, access tiers, and risk evaluations as part of the regulation on expanding secure access to Confidential Information Protection and Statistical Efficiency Act (CIPSEA) data assets.
1.7	OMB, in coordination with the ICSP, should promote the use of privacy-preserving technologies in the tiered access framework required under Title III of the Evidence Act by identifying an initial set of promising tools over the next 1 to 3 years.
1.8	OMB, in coordination with the ICSP, should identify models for shared responsibility among data providers and users and provide guidance on applying such models through the regulation on expanding secure access to CIPSEA data assets.
1.9	OMB should identify new frameworks for determining sanctions for unauthorized access, use, and disclosure of government data assets.
<b><i>The Standard Application Process</i></b>	
1.10	OMB should prioritize resources to optimize the user experience for the Standard Application Process.

*Table continues*

Rec. #	Recommendation
<b>Recommendations Part 2. Other Evidence Act Items</b>	
2.1	OMB should propose legislative text as part of the FY 2024 Budget to allow full implementation of Evidence Act Section 3575 for the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the Census Bureau.
2.2	The Federal Committee on Statistical Methodology should set a learning agenda promoting research on timely, relevant, and actionable questions, demonstrating the value and promise of the National Secure Data Service (NSDS).
<b>Recommendations Part 3. NSDS Functions</b>	
<i><b>Function 1. Coordination—Evidence Act Implementation for the Federal Government and Beyond</b></i>	
3.1	The NSDS should coordinate with the ICSP, communities of practice, and other key stakeholder groups to establish best practices for implementing the requirements of the Evidence Act and should serve as a model for testing and demonstrating those best practices for governments at all levels.
<i><b>Function 1. Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance</b></i>	
3.2	The NSDS website should serve as a “front door” to the nation’s data assets, organized around a set of personas that reflect basic user needs.
3.3	The NSDS should identify opportunities for automation of its “intake process,” providing a high-quality user experience while focusing staff effort on complex user needs.
3.4	The NSDS should employ data concierges to help users refine their research projects, discover relevant data, and acquire access to that data.
3.5	The NSDS should employ technical assistance leads who develop educational resources for data providers and data users related to the methods and technologies used by the NSDS and in the broader evidence-building ecosystem.
3.6	The NSDS should actively encourage the development of communities to crowdsource support for users, complementing the work of the data concierges.
3.7	The NSDS should regularly sponsor projects that demonstrate the value of streamlining data sharing and increasing coordination, specifically projects that highlight cross-functional, cross-agency, and cross-governmental topics.
<i><b>Function 2. Communication</b></i>	
3.8	OMB should adopt a clear statement of purpose for the NSDS that is rooted in its core value.
3.9	Through the NSDS website, the National Center for Science and Engineering Statistics (NCSES) should create a public information hub that serves as a central repository for resources about, and communications from, the NSDS.
3.10	NCSES, in coordination with OMB and the ICSP, should build a comprehensive communications strategy for the NSDS.

Table continues

Rec. #	Recommendation
3.11	The NSDS should build a system for routinely engaging with key partner groups and user communities for the purposes of needs sensing, operational improvement, and advocacy for the use of data to improve decisionmaking.
<b>Function 3. Research and Development</b>	
3.12	The NSDS should promote the use of privacy-preserving technologies that support working with data in situ, coordinating with the research community to develop efficient, scalable tools for users from all levels of government (including through open competitions).
<b>Function 4. Data Standardization</b>	
3.13	The NSDS should coordinate with stakeholders to develop and promote standards for government data at all levels.
<b>Recommendations Part 4. NSDS Organizational Structure and Governance</b>	
4.1	NCSES, in coordination with OMB and the ICSP, should leverage Congressionally appropriated demonstration project funding to establish the America’s DataHub Consortium (ADC) as the pilot foundation for the NSDS. The ADC should sponsor pilots that demonstrate how data service structures and functions could grow, adapt, and evolve over time to realize the Committee’s vision for the NSDS.
4.2	The NSDS should be a legally recognized entity that is owned by the federal government and operated by a contractor.
4.3	Through its governance and operations, the NSDS should model a holistic approach to transparency and accountability.
4.4	In its initial phases, NCSES, in coordination with the ICSP, should sponsor pilot projects that (1) explore the NSDS core functions—that is, coordination, communication, R&D, and data standardization, and (2) demonstrate the value of streamlining data sharing and coordination, specifically with projects that highlight cross-functional, cross-agency, and cross-governmental topics.
<b>Recommendations Part 5. NSDS Technical Infrastructure and Tools</b>	
5.1	The NSDS should provide technology so that users at any tier of access can safely and efficiently analyze data assets hosted by affiliated organizations, including federal, state, territorial, local, and tribal governments; nonprofits; and other organizations.
5.2	To support the discovery of data assets for evidence building, the NSDS should provide a technological process to support access to searchable and discoverable data, request data access, track the approval process, and document the outcomes of that process.
5.3	To support a seamless user experience, the NSDS should provide training and tools to harmonize the format and content across data inventories and catalogs and to ensure complete, consistent metadata are included in these inventories.
5.4	The NSDS should collect and house a searchable inventory of projects that highlights what data sets are being used for what purposes.

Table continues

Rec. #	Recommendation
5.5	The NSDS should provide tools and support to users in conducting secure, accurate, and scalable analyses. In addition, the NSDS should facilitate the development and deployment of data access protocols that offer alternatives to the standard direct data access models.
5.6	The NSDS should facilitate the development and application of statistical disclosure limitation methods.
5.7	To ensure auditability and transparency, the NSDS should provide tools to track the provenance of all data sets and all evidence-building analyses performed over those data sets.
5.8	The NSDS should provide tools and support to users in conducting scalable, privacy-preserving record linkages, facilitating data preparation and review of matching metrics. As part of its data concierge services, the NSDS should coordinate with federal, state, territorial, local, and tribal government officials seeking linkage services.
5.9	The NSDS should feature a sandbox for testing new and innovative technologies and software for multiple data access tiers, data protection protocols, and data analysis.
<b>Recommendations Part 6. Resources and Funding</b>	
<i>Resources to Enhance the Existing Evidence-Building Ecosystem</i>	
6.1	The OMB Director and Agency Heads, in consultation with the designated Chief Data Officer (CDO), Evaluation Officer, and Statistical Official at each agency, shall allocate funds from appropriations to adequately resource and support evidence-building activities for FY 2023. In addition, the OMB Director should prioritize direct appropriations and funding flexibilities as part of the FY 2024 Budget formulation process and encourage all Agency Heads to prioritize Evidence Act implementation activities going forward.
6.2	OMB, in coordination with the ICSP, the CDO Council, the Evaluation Officer Council, and other relevant federal councils, should develop a systematic approach to funding Evidence Act implementation.
6.3	The OMB Director should request funding and authority in the FY 2024 Budget for a new interagency budget account with transfer authority to support Evidence Act implementation.
6.4	The OMB Director should prioritize additional resources for OMB staff responsible for coordinating implementation of the Evidence Act Title III regulations, Title II guidance, Title I implementation activities, and other evidence-building priorities in the current fiscal year and in the FY 2024 Budget request.
6.5	The OMB Director should propose in the FY 2024 Budget request a new block grant for state, territorial, local, and tribal funding to support cross-program data infrastructure improvements and data modernization.
<i>NSDS Resources to Meet Its Mission</i>	
6.6	NSDS core functions should be funded through direct spending authority.
6.7	As data service functions and user demand for its services grow over time, the NSDS should explore a mixed funding model that leverages sustainable and dynamic funding approaches, including budget requests through NCSES, existing and new federal grant programs, repurposed agency funds, federal-state partnerships, private-sector support, a shared services model, and user fees for select services.

## 2. Background

This section provides background on the Evidence Commission; the Evidence Act; the evolving evidence-building ecosystem; and ACDEB’s purpose, progress, and promise.

### The Commission on Evidence-Based Policymaking

The Evidence Commission was created by the [Evidence-Based Policymaking Commission Act of 2016](#). With widespread bipartisan support, the law signaled the President’s and Congress’s commitment to building capacity to produce higher-quality evidence for decisionmaking. In its [2017 report](#), the Evidence Commission articulated a vision around data and evidence generation and use as a routine and critical function of government. The report features 22 recommendations to improve data access, modernize privacy and confidentiality protections, strengthen evidence-building capacity, and establish an NSDS to support governmentwide evidence building.

### The Foundations for Evidence-Based Policymaking Act

In 2018, Congress passed the [Evidence Act](#), addressing half of the Evidence Commission’s recommendations. The law established new legal expectations for openness and accessibility, building a framework where leaders across the government work together to coordinate data and evidence needs and uses. This includes aligning data from various sources, such as survey data and administrative data, with differing degrees of protection needed, from open to confidential, for different purposes, including producing statistical estimates and informing the administration and evaluation of government programs for certain shared purposes. Box 1 defines key terms and definitions.

## Box 1. Key Evidence Terms and Definitions

This box presents key evidence terms as defined by the Foundations for Evidence-Based Policymaking Act of 2018 and related guidance, as noted.

**Evidence.** Information produced as a result of statistical activities conducted for a statistical purpose. There are 4 categories of evidence, described by the Office and Management and Budget memorandum on “Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance” ([M-19-23](#)):

- **Foundational fact finding.** Foundational research and analysis such as aggregate indicators, exploratory studies, descriptive statistics, and basic research.
- **Performance measurement.** Ongoing, systematic tracking of information relevant to policies, strategies, programs, projects, goals/objectives, and/or activities.
- **Policy analysis.** Analysis of data, such as general purpose survey or program-specific data, to generate and inform policy, e.g., estimating regulatory impacts and other relevant effects.
- **Program evaluation.** Systematic analysis of a program, policy, organization, or component of these to assess effectiveness and efficiency.

**Statistical activity:** (1) The collection, compilation, processing, or analysis of data for the purpose of describing or making estimates concerning the whole, or relevant groups or components within, the economy, society, or the natural environment; and (2) includes the development of methods or resources that support those activities, such as measurement methods, models, statistical classifications, or sampling frames.

**Statistical purpose:** (1) The description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups; and (2) includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support these purposes.

Congress’s intent for prioritizing evidence building is woven throughout the Evidence Act, including the following:

- The Evidence Act’s Title I requires multi-year learning agendas, or evidence-building plans. In addition, Title I includes requirements for analyzing the capacity of federal agencies to engage in evidence-building activities. Agency “capacity assessments” assess agency capacity to support the development and use of evaluation.
- Title II, or the OPEN Government Data Act, establishes that the default for government data is open availability unless otherwise prohibited by law and requires agencies to publish data inventories. Notably, this requirement applies broadly to government data assets to support transparency and has implications and foundational benefits for evidence building across government.

- Title III, or the Confidential Information Protection and Statistical Efficiency Act of 2018 (CIPSEA), requires federal agencies to make data accessible to federal statistical agencies within a strong privacy framework and for those statistical agencies to ensure secure access to confidential data assets and to create a Standard Application Process (SAP).

Collectively, these provisions indicate that the Evidence Act intends to increase the accessibility of data for evidence building. While the CIPSEA provisions give responsibilities to statistical agencies that require them to operate as a federal statistical system, the Evidence Act is not exclusively about improving the statistical system or statistical data but aims to improve the whole of government.

## **Evidence Act's Title I and the Role of Evaluation Officers and Statistical Officials**

The Evidence Act requires systematic planning through learning agendas, capacity assessments, and evaluation plans. Newly established Evaluation Officers coordinate this planning in collaboration with other newly established officials—that is, Chief Data Officers (CDOs) and Statistical Officials. Evaluation Officers oversee evaluation activities, including providing technical and methodological resources to assess, improve, and advise on evaluations and conceptualizing, prioritizing, and designing related activities. They are also charged with encouraging evaluation activities within their respective agencies to build evaluation capacity over time. For more information about these activities, see [evaluation.gov](https://www.evaluation.gov).

Statistical Officials support these activities by advising on statistical policy, techniques, and procedures with a focus on championing data quality and ensuring data confidentiality. For more information on the role of CDOs, see Evidence Act's Title II (OPEN Government Data Act) and the Role of Chief Data Officers below. OMB [M-19-23](#) provides more information on these three roles.

## **Evidence Act's Title II (OPEN Government Data Act) and the Role of Chief Data Officers**

The Evidence Act establishes a new capacity for data management and governance by creating CDOs. Under the law, these officials have an extensive set of responsibilities, which are delegated in some instances to the heads of statistical agencies and units when needed to comply with statistical law (including CIPSEA).

As required by law, CDOs across the federal government are establishing Data Governance Boards in their agencies and are coordinating the participation of Evaluation Officers, Statistical Officials, and others with these boards. In addition, CDOs are producing data inventories and are formulating a community of practice in the [CDO Council](#). The OPEN Data Government Act also requires OMB to issue guidance for implementing Title II of the Evidence Act, some of which has not been issued as of the date of this report.

## **Evidence Act’s Title III (CIPSEA) and the Role of the Federal Statistical System**

The Evidence Act provides new authority and responsibilities to federal statistical agencies. Specifically, Congress granted responsibilities to statistical agencies that the Evidence Commission initially envisioned for the NSDS (i.e., the Presumption of Accessibility, the responsibility to expand secure access to federal data assets, and the SAP). This authority strengthened agencies as trusted intermediaries between data providers and data users without creating a new entity. The NSDS must fit within this existing framework.

Federal statistical agencies have a clear mission to produce data and evidence and a long history of delivering value to the American public. This provides a strong foundation and core infrastructure for advancing evidence-based decisionmaking, especially since statistical agencies have established and trusted relationships with state, territorial, local, and tribal data providers that are essential to unleashing the potential of administrative data sets held at those levels.

Title III of the Evidence Act establishes a framework to expand agency capacity and enable access to the federal government’s confidential data for evidence building (that is, provide value) while protecting those data. Table 1 describes the regulations required by CIPSEA 2018 and outlines expected roles and responsibilities for OMB, federal statistical agencies, and other federal agencies as stated or implied in statute.

**Table 1. CIPSEA 2018 Regulations**

Section	Requirements, Roles, and Responsibilities
<b>Responsibilities for statistical agencies (Sec. 3563)</b>	<p><b>Statistical agencies:</b> In general, carry out the following primary responsibilities:</p> <ul style="list-style-type: none"> <li>• Produce and disseminate relevant and timely information,</li> <li>• Conduct credible and accurate statistical activities,</li> <li>• Conduct objective statistical activities, and</li> <li>• Protect the trust of information providers by ensuring the confidentiality and exclusive statistical use of their responses.</li> </ul> <p><b>All federal agencies:</b> Enable, support, and facilitate statistical agencies in carrying out responsibilities.</p> <p><b>OMB:</b> Prescribe appropriate regulations.</p>
<b>Presumption of Accessibility for Statistical Agencies and Units (Sec. 3581)</b>	<p><b>OMB:</b> Prescribe appropriate regulations, require timely provision of data assets, provide a list of statutes that exempt agencies, establish clear and consistent standards for complying with the Privacy Act or other applicable laws, and require a transparent process for statistical agencies to request data assets and for agencies to respond.</p> <p><b>All federal agencies:</b> Make any data asset available to any statistical agency or unit for evidence building unless prohibited by statute; comply with the regulation.</p>
<b>Expanding Secure Access to CIPSEA Data Assets (Sec. 3582)</b>	<p><b>Statistical agencies:</b> Expand access to data assets acquired or accessed to develop evidence while protecting against inappropriate access and use; comply with the regulation.</p> <p><b>OMB:</b> Issue a regulation that addresses the following:</p> <ul style="list-style-type: none"> <li>• Standards assessing each data asset owned or acquired by the statistical agency or unit to categorize the sensitivity level and assign the corresponding level of accessibility;</li> <li>• Standards to improve access to a data asset, including methods to minimize disclosure risk for data subjects; and</li> <li>• A requirement for each statistical agency or unit to conduct a comprehensive risk assessment of data prior to public release, including standards for such risk assessments and criteria for release decisions.</li> </ul>
<b>Standard Application Process (SAP) (Sec. 3583)</b>	<p><b>Statistical agencies:</b> Implement the process; comply with the regulation.</p> <p><b>OMB:</b> Consult with stakeholders, including the public, agencies, state and local governments, and non-governmental researchers; establish a process for parties (including state, local, and tribal governments) to apply for access to data assets accessed or acquired by statistical agencies for evidence building, including standards for full transparency.</p> <p>Major milestones include the following:</p> <ul style="list-style-type: none"> <li>• <a href="#">Request for public comments</a> on the SAP policy released in January 2022</li> <li>• SAP data catalog posted through <a href="#">ResearchDataGov</a> in spring 2022</li> <li>• Expected web portal launch as the single “front door” for SAP in late 2022</li> </ul>

Note. While OMB continues to confirm this as a priority, the regulations have yet to be issued as of the date of this report.

CIPSEA Confidential Information Protection and Statistical Efficiency Act  
 OMB Office of Management and Budget  
 SAP Standard Application Process

# The Evolving Data and Evidence Ecosystem

## The Actors and the Infrastructure

The U.S. evidence-building ecosystem is broad and highly decentralized, encompassing the 13 federal principal statistical agencies, over 100 federal agencies or units that engage in statistical activities, federal programmatic agencies, tens of thousands of state and local government entities that provide administrative data for statistical purposes and use evidence for their own decisionmaking, private-sector data providers and users, and the public.

In addition to the federal statistical agencies, Statistical Officials, Chief Data Officers, and Evaluation Officers are explicit statutory actors in the ecosystem, designated to coordinate and connect parts of the ecosystem to each other. Here, again, the NSDS must fit within this framework.

There are many existing and emerging entities building evidence and providing services for data providers and users with objectives complementary to what is envisioned for the NSDS. Among the numerous federal, state, and private-sector data infrastructures and evidence-building programs, the Committee examined the following:

- [America's DataHub Consortium](#) (ADC), sponsored by the National Center for Science and Engineering Statistics (NCSES) at the National Science Foundation (NSF), engages in activities that facilitate data access and sharing, security and infrastructure, and analysis with a science and engineering focus. The ADC leverages NSF's funding flexibilities to solicit project solutions from consortium members, including nonprofit entities, small and large businesses, and academic organizations, and to make awards.
- [COVID-19 Research Database](#) is a research platform that harnesses a vast amount of health data to enable policymakers and researchers to conduct robust research on the effects of the pandemic while preserving privacy. A coalition of industry leaders, including technical and data partners, came together to develop the database in just 6 weeks.
- [Federal Statistical Research Data Centers](#) (FSRDCs), operated by the U.S. Census Bureau in partnership with federal statistical agencies and research institutions, provide secure facilities throughout the United States that enable authorized individuals to access restricted-use microdata for statistical purposes.
- The [Midwest Collaborative](#), a coalition of state workforce and education agencies, is collaborating with the Coleridge Initiative and regional university partners to design a system that enables individual states to answer critical questions that are relevant to societal well-being. State partners set priorities, contribute data, and establish research and data product agendas. The Coleridge Initiative provides training and a federally approved and secure cloud-based computing platform. Founding states include Iowa, Illinois, Indiana, Kentucky, Michigan, Missouri, New Jersey, Ohio, and Tennessee. Key products that demonstrate the possibilities of cross-state collaboration include the [Multi-State Postsecondary Dashboard](#) and the [Unemployment to Reemployment](#) portal.

- The [National Artificial Intelligence Research Resource](#) is envisioned as a shared computing and data infrastructure that will provide AI researchers and students across scientific fields and disciplines with access to computational resources and high-quality data, along with appropriate educational tools and user support.
- The [State Wage Interchange System](#) is an agreement among states to share employment insurance wage record data so that workforce agencies can track individuals who participated in [Workforce Innovation and Opportunity Act](#) (WIOA) activities but then moved to another state. The program is primarily for reporting required outcome measures as part of WIOA and makes use of the Unemployment Insurance Interstate Connection Network. While this is a single-purpose data-matching facility, the states and their federal partners at the Department of Labor and the Department of Education are exploring possibilities that could make data available for broader evidence-building activities.

For other examples, see the report appendices and the Supplemental Information [posted](#) with this report.

## The Evolution

The evidence ecosystem—and the work to support it—continues to evolve. Many groups across the public and private sectors, like those described above and presented in the appendices and Supplemental Information for this report, are focused on improving data access, linkage, and analysis to support better decisionmaking. New programs, initiatives, and infrastructures are also joining the ecosystem at an accelerating pace.

With the ecosystem evolving rapidly, it is important to focus on opportunities that exist right now to lay the foundation for the future. This includes progress on Evidence Act implementation, a focus on the role of state and local stakeholders, advances in privacy-preserving technologies, and developments on the federal legislative and budget fronts.

***Federal statistical system implementation of the Evidence Act.*** The Evidence Act has the capability to facilitate the evolution of the federal statistical system if the full force of the statutory provisions is used by OMB and federal agencies. Federal statistical agencies are implementing the Evidence Act provisions, as required by law. The Interagency Council on Statistical Policy (ICSP) launched the SAP through a single portal that features a data catalog to guide users to data assets for their evidence-building projects. In addition, OMB is developing the required CIPSEA 2018 regulations and guidance, in partnership with the ICSP to include the views and perspectives of the federal statistical system. These regulations should accelerate progress in implementing the functions that the Evidence Commission envisioned for the NSDS and should form the basis of many of the recommendations of this Committee.

***State and local stakeholders.*** State, territorial, local, and tribal governments are critical actors in the evidence-building ecosystem—both as providers of data and evidence and as users of evidence for decisionmaking. Their varied roles and capacity needs, however, differ in important ways from those of federal agencies. Regular engagement among government stakeholders at all levels is critical for advancing the evidence ecosystem, including expanding Evidence Act frameworks to accommodate needs at the state and local levels.

The Evidence Act lays out requirements for federal agencies; however, its principles, frameworks, and standards have the power to advance evidence-based decisionmaking across the broader ecosystem. For instance, CIPSEA standard data sensitivity levels, tiered access frameworks, and comprehensive risk assessments can serve as models for state and local governments seeking to harness data for evidence building and aid interoperability across systems and data sets. Federal programmatic agencies can provide clarity around legal frameworks and tools, such as data-sharing protocols, for data collected by states as part of administering federal programs. The federal government can also provide a sustainable stream of resources that facilitates federal-state information sharing and communication.

The experiences of state, territorial, local, and tribal governments have great potential for informing evidence building at the federal level and for building products that can be used to improve the administration of federal programs. As states increasingly use their own data to inform their decisionmaking, there is a substantial opportunity for federal officials to use these state data to evaluate federal programs when permitted by state law. State agencies can also provide critical information products for the diverse range of users across their local regions that complement the national statistics produced by federal agencies. For more information, see Recommendations—Introduction: Foundations and Opportunities.

***Privacy-preserving technologies.*** New methods to securely link, analyze, and publish data will continue to revolutionize evidence-building efforts. Approaches including synthetic data with validation servers and secure remote access are used by federal statistical agencies today. Active research programs in government, academia, and the private sector are making federated data joins, secure multiparty computation, and new algorithms to make protecting data outputs efficient and practical. The Evidence Commission recommended that the NSDS coordinate research and development on and facilitate adoption of these methods.

***Federal legislative and budget developments.*** There are developments at the federal level that highlight the importance of evidence-based policymaking and signal the federal government’s commitment to evidence-building activities.

In July 2022, Congress passed the [CHIPS and Science Act](#), establishing an NSDS demonstration project under NSF’s NCSES and authorizing funding for each of the fiscal years 2023–2027. Previously, the House of Representatives had twice passed legislation that would authorize the NSDS. This Act signals strong support from Congress for the principles of the Evidence Act and the importance of the NSDS.

The fiscal year 2023 President’s Budget promotes evidence-based decisionmaking in federal agencies and outlines a blueprint for the future of the federal statistical system. This blueprint builds on the statutory framework provided by CIPSEA 2018 and depends on statistical agencies operating as a seamless system. The [budget](#) spotlights proposed evidence-building initiatives of the federal statistical agencies, including funding at NCSES to “lead Government-wide development of evidence-building infrastructure such as the Standard Application Process, America’s Datahub Consortium, and early work on a National Secure Data Service.”

Throughout the report, the Committee highlights interdependencies and opportunities across the evidence ecosystem that have the potential to advance the data evolution. For more information, see Recommendations—Introduction: Foundations and Opportunities, Recommendations Part 4. NSDS Organizational Structure and Governance—America’s DataHub Consortium as the Foundation for the NSDS, the report appendices, and the Supplemental Information [posted](#) with this report.

## ACDEB’s Purpose, Progress, and Promise

### Purpose

ACDEB is charged with providing recommendations on how to promote the use of federal data for evidence building. This includes advising OMB on forthcoming CIPSEA regulations and other Evidence Act items and on the value of the NSDS within the evidence ecosystem.

### Progress—the Here and Now

As described above, the CIPSEA 2018 requirements are designed to enhance how the statistical system facilitates evidence building, just as the Evidence Commission envisioned, and even before the NSDS exists. At ACDEB’s January 2022 meeting, Commissioner of the Bureau of Labor Statistics and Chair of the ICSP, Bill Beach, described this connection:

***“We [that is, the federal statistical system] do not see implementing CIPSEA 2018 and building a data service as two separate activities. They are heavily overlapping, and we are, in effect, building parts of the data service now.”***

In the Year 1 report, the Committee recommended that the OMB Director take immediate steps to issue guidance and regulations required under the Evidence Act that provide necessary frameworks to inform the development of the NSDS.

Throughout its second year, ACDEB engaged with OMB and the ICSP on the CIPSEA 2018 guidance and regulations. This engagement offered the Committee the opportunity (1) to provide feedback on the development of the CIPSEA 2018 regulations and guidance and associated implementation activities, and (2) to use these efforts to inform the Committee’s work. These conversations provided key steppingstones to connect current realities to the target vision for the NSDS. The Committee’s recommendations include advice on the following aspects of the Evidence Act:

- Primary responsibilities of statistical agencies and trust,
- Presumption of Accessibility,
- Expanding secure access to CIPSEA data assets,
- Standard Application Process,
- Sharing of business data among designated statistical agencies,
- Federal statistical system-wide methodological learning agenda, including privacy-preserving technologies, and
- How the NSDS would fit within existing legal and statistical frameworks.

For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance and Recommendations Part 2. Other Evidence Act Items.

## Promise—the Vision for the National Secure Data Service

The ACDEB Year 1 report recognized the need for and value of a potential NSDS “to facilitate data access, enable data linkages, and develop privacy-enhancing techniques in support of increasing data for evidence building across the entire evidence-building ecosystem.” Specifically, the Year 1 report highlighted the following:

- **Value.** The NSDS will produce value for the American public by facilitating evidence building. Federal agencies, with the NSDS as a coordinator, will operate as a unified evidence-building system, partnering with state and local governments and nongovernmental organizations.
- **Structural requirements.** The NSDS will be a legally recognized entity with hardware, software, and administrative infrastructure and capacity.
- **Required organizational attributes.** The NSDS will exhibit the following attributes: (1) transparency and trust; (2) accessibility; (3) independence; (4) legal responsibility to acquire, protect, and link data; (5) scalable functionality; (6) sustainability; (7) oversight and accountability; and (8) intergovernmental support.
- **Core functions.** The NSDS will do the following:
  - **Coordinate.** Coordinate and support evidence-building efforts that cut across entities by facilitating linkage of, secure access to, and analysis of nonpublic data and by providing capacity-building services for data users, data providers, and related communities of practice.
  - **Communicate.** Communicate the value and use of data for evidence building and how the data are protected.
  - **Research and development (R&D).** Facilitate R&D and adoption of practices and methods that enhance privacy and confidentiality and improve record linkage quality.
  - **Data standardization.** Foster and promote data standardization to enable more efficient and high-quality linkage, access, and analysis.

Throughout its second year, the Committee considered in more depth how the NSDS would enter the existing evidence ecosystem as a new partner, striving to complement and enhance—rather than supplant or replace—current actors, infrastructures, and capabilities. The NSDS should help fill gaps for other actors in the ecosystem through its core functions. For more information, see Recommendations Part 3. NSDS Functions. Furthermore, the NSDS should support a broader capacity-building strategy around evidence building to spur new ideas and to create value. For more information, see Recommendations Part 6. Resources and Funding.

The Committee expects the National Secure Data Service to benefit the American public by connecting *federal, state, and local partners* with the services *they* need to securely share, link, and analyze data *to support* the use of evidence in policymaking.

## Acknowledgments

The Office of Management and Budget (OMB) delegated responsibility for administering the Committee to the Office of the Under Secretary for Economic Affairs (OUSEA) at the Department of Commerce and, by extension, the Bureau of Economic Analysis (BEA) and the Census Bureau. In addition, the Committee received ongoing support from the National Center for Science and Engineering Statistics (NCSES) and OMB. The Year 2 report offers a window into this team’s tireless efforts and commitment to excellence in supporting the Committee. The Committee would like to express its gratitude to the following individuals who provided invaluable assistance during the Committee’s second year:

Avi Alpert, AA Strategy Consulting

Jeannine Aversa, BEA

Jessica Comola, PalTech

Peter Fisk, BEA

Danielle Helta, BEA

Alyssa Holdren, BEA

Francise Jackson, NCSES

Colby Johnson, BEA

Gianna Marrone, BEA

Shelly Martinez, OMB

David Mendez, BEA

Kenneth Pond, BEA

Elizabeth “Meagan” Tydings, Census Bureau

Ryan Wist, OUSEA

# 3. Recommendations

## Introduction: Foundations and Opportunities

For decades, the use of data, statistics, evidence, and evaluation has been advancing. These activities, however, have not always been coordinated or resourced at appropriate levels. The federal statistical system, the evaluation community, Chief Data Officers, administrative programs, state and local governments, and many others have made tremendous contributions, but there is considerable room for improvement, learning, innovation, and continued progress. Indeed, the evidence ecosystem is primed to demonstrate the value of using data and evidence to learn and adapt, including applying lessons about what works, when, and where.

This section describes some of the foundations and opportunities for evidence building from the experiences of federal agencies—both statistical and programmatic—and governments at other levels.

### Federal Statistical Agency Experience

The strong foundations of the statistical agencies must inform the implementation of the Evidence Act, and the Evidence Act must inform their operations.

**Foundations.** Statistical agencies have a clear mission to produce data and provide access to evidence. To meet this mission, agencies must address the explicit tradeoff between utility and privacy. Agencies have extensive experience applying statistical disclosure limitation methods to produce data sets for public consumption. Statistical agencies also have decades of experience providing access to confidential data through secure enclaves, including Federal Statistical Research Data Centers. Furthermore, federal statistical agencies are leading the way to develop and deploy new methods that increase data availability, utility, fairness, and equity while protecting privacy and confidentiality. The federal statistical system is poised to apply such methods more widely.

**Opportunities.** Since most laws do not explicitly address the use of data for evidence building, there is often a high burden to demonstrate that such use is allowable. Through the requirements to expand access to CIPSEA data assets, the Evidence Act demands transparent approaches to determine sensitivity levels, access tiers, and output controls. In addition, many CIPSEA-designated statistical agencies and units rely on sample surveys as a primary data source and prioritize mandatory or other sought-after public products. The Evidence Act’s Presumption of Accessibility provides a clear authority for statistical agencies to acquire federal administrative data and make those data available for evidence building. Statistical agencies, especially smaller statistical units, have workforce and capacity constraints, which lead to difficulties in supporting evidence building for diverse user groups.

## Federal Programmatic Agency Experience

The experience of federal programmatic agencies (e.g., the Department of Housing and Urban Development and the Department of Veterans Affairs) can inform the implementation of the Evidence Act, and the Evidence Act can inform the operations of these nonstatistical agencies by defining the value of diverse types of evidence and making data more accessible to improve programs.

**Foundations.** Federal programmatic agencies have a mission to deliver effective government programs. Many agencies have invested in technology improvements to identify waste, fraud, and abuse. Furthermore, these agencies provide relevant information, such as societal trends and benchmarks for performance, to inform programs and policies.

**Opportunities.** While the Evidence Act formalized new roles and authorities of evidence producers and enablers, including charging agencies with identifying Chief Data Officers, Evaluation Officers, and Statistical Officials, the law’s power to transform the capacity of data and evidence to improve programmatic delivery and create value has not been fully realized. Program agencies have large data stores that could substantially inform decisionmaking, but they are currently inaccessible for broad evidence-building purposes. By implementing the Evidence Act provisions on the Presumption of Accessibility and expanding secure access to federal data assets, agencies can unleash these data for evidence building. There is also limited ability and desire to build common standards across departments and limited data inventories that provide transparent information on what data are available, what topics have been studied, with what methods, and by which experts.

## State, Territorial, Local, and Tribal Government Experience

The experience of state and local government agencies can inform the implementation of the Evidence Act, and the Evidence Act can inform the operations of state and local governments.

**Foundations.** State and local governments primarily deliver services to the public. Through their programmatic and funding relationships, state and local government agencies have a long history of working closely with federal agencies. Federal agency activities, including program delivery and evaluation, rely on these federal-state partnerships to produce data and create evidence for many public policy areas, including health, labor, education, and justice. Likewise, a significant amount of data managed by the federal government is gathered and reported by state and local government agencies and their partners, making such agencies key to ensuring quality data for timely, local, and actionable decisionmaking.

**Opportunities.** Although administrative data held by state and local governments offer much promise for improving evidence-based decisionmaking, these records are typically of varying quality, often poorly documented, differ in structure across states, and frequently lack robust identifiers for vulnerable and marginalized populations because they are not collected for evidence-building purposes. In addition, the lack of clear legal interpretations often inhibits the wider use of data among agencies and across state lines. State, territorial, local, and tribal governments necessarily focus workforce and technical capacity on running programs. They are not well-resourced to do their own evidence building and rarely have the capacity to support others’ needs for their data.

Furthermore, the appetite and ability for states to engage in greater data sharing and linking also vary substantially across the country. Some states have legal frameworks that make it possible to share data with federal agencies or private intermediaries, while others would not voluntarily give their data to the federal government. Likewise, some states and localities have robust data linkage programs, but many jurisdictions lack this capability. State and local decisionmakers could also benefit from more systematic efforts to share effective policies and practices with officials in other jurisdictions who are running similar programs.

In addition, there is a substantial opportunity for federal officials to use state data to evaluate programs that are delivered at the state and local levels. State administrative data could provide valuable insights to inform federal decisionmaking, and federal agencies could adopt best practices modeled by the states. State data are often more specific to small geographies and more current than federal data. As such, state data could prove invaluable for monitoring the impact of government interventions intended to benefit small regions recovering from economic or natural disasters in a timely and actionable manner.

The Committee’s recommendations and the supporting evidence and findings are presented below. The organization of the recommendations mirrors the Committee’s charge—starting with recommendations on the “here and now” (that is, the forthcoming Evidence Act regulations and other Evidence Act items) and flowing through to the vision for the NSDS, including its functions, organizational structure, governance, and technical infrastructure, and how the evidence ecosystem should be supported through a comprehensive funding strategy.

## Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance

**Recommendation 1.1. The OMB Director, in coordination with the ICSP, should issue guidance and regulations required under the Evidence Act, incorporating feedback from ACDEB’s Year 1 report, subsequent engagements with the Committee, and this report.**

The Year 1 report stated: “The OMB Director should take immediate steps to promulgate draft guidance and regulations required under the Evidence Act.” While many of these regulations are still outstanding, through its second-year discussions with OMB and the ICSP, the Committee provided advice on how to move forward on portions of this recommendation.

The Committee reaffirms the importance of the required Evidence Act regulations that provide the necessary framework to inform the development of the NSDS and to support evidence-based policymaking throughout the federal government. OMB’s guidance should aim to increase consistency, timeliness, and predictability of access decisions and should substantially increase access by employing best practices and the best technology for privacy protection.

With a broad vision for the evolving evidence ecosystem in mind, the Committee recommends the following:

**Recommendation 1.2. OMB, in coordination with the ICSP and other federal councils, should promulgate required Evidence Act regulations and guidance with an eye toward supporting the NSDS functions, as envisioned by the Committee.**

The Committee notes that the NSDS must operate under these regulations and recognizes that recommendations in subsequent sections of the report directed toward the NSDS may apply to the broader federal statistical system. As such, OMB should consider the totality of the report when issuing the forthcoming CIPSEA regulations and guidance and setting policies that affect the NSDS. Given the intention of the Evidence Act to increase the availability of federal data for evidence building, the CIPSEA guidance should be specific enough to ensure greater compliance and streamline access to data.

The next section provides recommendations on the CIPSEA regulations, aimed at meeting the requirements of the Evidence Act and positioning the evidence-building ecosystem to function more effectively. The Committee provides advice on both the content of the regulations and how the regulations should be implemented.

## Primary Responsibilities of Statistical Agencies and Trust

### *Regulation Content and Implementation Strategy: Promote Public Trust*

**Recommendation 1.3. OMB should apply practical principles for promoting public trust and should empower all federal agencies to participate in this process.**

As OMB develops and considers publication of its regulation on public trust under the Evidence Act, it must take into consideration broad factors and principles about how to foster and encourage public trust in data. The Committee urges OMB to empower all agencies to participate in this process, not just the federal statistical agencies. In addition, the Committee encourages OMB and the statistical agencies to specifically apply practical principles for promoting public trust in the future that align with historical practices, such as managing the risks of disclosure, planning for future data users, communicating effectively and appropriately with the American public about data uses, and applying tiered access frameworks that recognize the availability for more data, not less, within strong privacy frameworks.

The success of the evidence ecosystem hinges on public trust in the institutions and organizations collecting the data and providing the infrastructure for data access and analysis. A primary responsibility of statistical agencies, as codified by the Evidence Act, is “to protect the trust of information providers by ensuring the confidentiality and exclusive statistical use.” The federal statistical system provides core infrastructure for the evidence ecosystem with its routine surveys and data collections. These statistical instruments yield robust and reliable information about the country and economy. In other words, statistical agencies foster trust by creating value for the taxpayer through the data they collect.

The statistical agencies also serve as a model for other agencies in government and other data users in the appropriate approaches for protecting and sharing data, institutional cultures for data governance and protection, and capabilities for promoting data innovation and use. Even while recognizing this strong history, the Committee acknowledges there is a continuous need for learning and institutional growth in the years ahead to sustain and even improve public trust.

## The Presumption of Accessibility for Statistical Agencies and Units

### *Regulation Content and Implementation Strategy: Maximize the Impact of Federal Administrative Data for Evidence-Based Decisionmaking*

#### **Recommendation 1.4. OMB should issue the Presumption of Accessibility rule to maximize the impact of federal administrative data for evidence-based decisionmaking.**

The Presumption of Accessibility creates a “yes, unless” standard for federal agencies to share administrative data with statistical agencies for the purpose of evidence building. The ICSP, in coordination with federal programmatic agencies, state and local data providers, and other key stakeholders, must: (1) establish structures so that administrators of federal, state, territorial, local, and tribal governments can use federal data assets to help improve the administration of their own programs without revealing the identity of any individual entity in the data and (2) facilitate the use of these data for evaluation and analytics while reducing reporting burdens.

In the Year 1 report, the Committee acknowledged that “...within the broader evidence-building ecosystem in the United States, there are thousands of state and local governments who collect administrative data that have vast potential for informing evidence-based decisionmaking. Often, these data are gathered through federally funded or run programs and thus are, in part, being shared with the federal government today; however, the power of administrative data for evidence building has not yet been realized.”

Through its Year 2 discussions with OMB and the ICSP, the Committee explored applications of the Presumption of Accessibility, including providing federal statistical agencies with secure access to administrative data from other federal agencies: (1) to improve the core statistical products and (2) to be made available to meet evidence-building needs more widely.

Committee members noted many barriers to the use of administrative data, including the following:

- State and local program administrators spend a large amount of time and resources meeting the different reporting requirements of the various federal programs in which they participate. This often leads to duplication of efforts and does not leave the bandwidth for state and local officials to explore better ways to leverage data collected through these programs for their own decisionmaking.
- There are distinct legal and regulatory regimes that complicate data management. For example, state data shared with the Census Bureau for the [Longitudinal Employer-Household Dynamics](#) (LEHD) program are backed by individual data-sharing agreements with each state. Using the data for a new purpose involves reviewing and possibly renegotiating the various agreements.
- Administrative data often lack the documentation and quality to make them useful for evidence building. Agencies collecting and managing the data lack clarity from OMB that federal funds can support data capacity and management efforts.

The Committee also discussed ways in which a better two-way data and evidence flow could improve decisionmaking across all levels of government, noting that “[administrative] data regularly flow ‘up’ to the federal government but do not flow back ‘down’ to the data providers for their own evidence-based decisionmaking.” The National Center for Health Statistics (NCHS) [National Vital Statistics System modernization](#) efforts and the Department of Labor’s [Equity Data Partnerships](#) demonstrate the possibilities for leveraging existing federal-state relationships to improve state infrastructure, develop new products, and provide timely and actionable evidence for decision-making and program improvement. For more information, see the Committee’s use case reports on health and labor market activity in the Supplemental Information [posted](#) with this report.

In addition to the challenges states have with using the data they collect to administer federal programs for their own decisionmaking, states often face difficulties sharing data within their own jurisdictions. For more information, see Box 2. South Carolina Data Sharing Example.

## Box 2. South Carolina Data Sharing Example

South Carolina is actively advancing sound policymaking and fiduciary decisions based on evidence, so much so, that the General Assembly created, by law, a [Coordinating Council for Workforce Development](#) (CCWD).

The purpose of the CCWD is to engage in discussions, collaboration, and information sharing to support the state’s ability to prepare and train workers to meet current and future workforce needs. CCWD member agencies collect various types of information. The ability to share that information depends on the type and sensitivity of the data collected. Much of the data is publicly available, but it can be difficult to access those data because they are not stored and maintained in a central location.

More data linkages would assist the CCWD in providing evidence-based recommendations to improve workforce outcomes for all citizens and businesses in South Carolina. By eliminating personal details and making data available on an aggregate and anonymized basis, the CCWD would be able to fulfill statutory mandates to make recommendations to the General Assembly regarding programmatic improvements and other matters that the individual agencies do not currently have the authority to implement and for which legislation is required.

South Carolina’s [health and demographics integrated data system](#) is one of the most comprehensive and “mature” state-level systems in the country; however, efforts to incorporate additional data streams have struggled. Specifically, efforts to address and improve interagency data sharing related to economic development, workforce development, and education have been met with some resistance, sometimes small and manageable, other times not. Challenges to better data sharing include the following:

- Agencies are protective of their data because of the common fear that data will be used against them by the legislature or the media. Most state agencies do not like to see their names in the press as they receive very few positive stories. Especially for regulatory agencies, the data may be less than flattering and only reinforce narratives about the agencies.
- Another barrier is that the data being requested may not be the “right” data to answer a question. In this case, by limiting data access, agencies are being less “protective” in order to be better stewards of their data. Clarity about the request and the desired outcome often solves this problem.
- The greatest barrier may be how the government will use longitudinal data, and the state has yet to overcome this barrier. This is a cultural issue more than a process issue as the public fears what the government knows about them and how personal data would be used. Legislation in South Carolina to improve longitudinal data efforts failed because of public outcry. There is a pervasive concern that the government is “selling” citizens’ data or trying to force a student into a job because of something in the data. In addition, recent situations such as Facebook and data hacks have created a negative perception of the government using citizens’ data against them. Concepts such as de-identified data, data security plans, and oversight seem to do little to strengthen the argument for data sharing.

Success exists in the efforts to make data sharing a more standard operation within government, but large-scale legislative action will not be a means to that goal in the foreseeable future. The potential benefits of data sharing are invaluable, and data sharing requires trust.

The Committee explored innovative models like the Midwest Collaborative, through which states partner to create products and programs. States combine administrative data sets using standardized models within a secure computing platform, in pursuit of better policymaking and improved program administration. For more information, see ACDEB’s use case reports on education and workforce and labor market activity in the Supplemental Information [posted](#) with this report.

State and local governments could build on the principles of the Presumption of Accessibility and leverage new mechanisms, like the Midwest Collaborative, to expand the use of administrative data for evidence building. As a coordinator across the evidence-building ecosystem, the NSDS would be well-positioned to facilitate these activities. For more information, see Recommendations Part 3. NSDS Functions—Coordination.

### ***Implementation Strategy: Data-Sharing Agreements Across Federal Agencies***

**Recommendation 1.5. OMB, in coordination with the ICSP and other relevant federal councils, should identify mechanisms for streamlining data-sharing agreements across federal agencies.**

A major challenge in implementing the Presumption of Accessibility is the lack of clear guidance around how Memoranda of Understanding (MOUs) are developed, approved, and maintained to permit data sharing across federal agencies and with non-federal entities. Overcoming this challenge is consistent with the intent of the Presumption of Accessibility to expand the use of federal administrative data for evidence building.

Streamlining the MOU process could include developing standard MOUs, clarifying the legal framework for sharing data across states and agencies, expanding the use of existing templates, and establishing best practices, including encouraging multi-year agreements that anticipate recurring needs based on the same justification. This work should be a vehicle for identifying legal and policy impediments to data sharing and for developing a consistent data-sharing approach. The resulting findings should be used as a basis for recommendations on changes in law and regulations.

In the Committee’s Year 1 report, the Legislation and Regulations focus area recommended that “The OMB Director, working in consultation with the Interagency Council on Statistical Policy, the Chief Data Officers Council, and the Federal Privacy Council, should facilitate the creation of updated model language for MOUs (or a shared standard for data sharing) to facilitate data sharing and linkage projects and reduce the number of MOUs required to conduct regular evaluations of programs and policies.”

As part of its second-year investigative process, the Committee discussed many obstacles surrounding data-sharing agreements, including administrative challenges to managing the sheer number of MOUs, the need to renegotiate agreements over time, and differences in legal interpretations that inhibit data sharing.

The Committee also explored ways to overcome legal hurdles. One such example is the [Advanced Education Research and Development Fund](#) (AERDF), supported by the [DataSafes](#) program. The goal of the AERDF is to provide a secure environment for partnering with school districts to contribute data and to build analyses, conduct research, and generate meaningful insights. A key to facilitating this effort is the National Research Data Privacy Agreement, a standardized agreement developed by the [Student Data Privacy Consortium](#) that creates a contract between schools and researchers on the use of student data for research purposes and includes compliance procedures to help ensure privacy, confidentiality, and security.

Furthermore, the Committee discussed the broader philosophy of the DataSafes program. As stated by Wade Shen during a meeting with the Technical Infrastructure subcommittee: “The law and technology must co-evolve.” Technology demonstration, and the risk assessment associated with it, should inform agency policies. For more information, see [Expanding Secure Access to CIPSEA Data Assets](#) below and [Recommendations Part 3. NSDS Functions—Research and Development](#).

In addition, as highlighted by the Evidence Commission, “in 2014, OMB developed an optional model agreement for agencies to consider, though uptake has been limited to date.” The opportunities for developing a standard MOU template and multi-year agreements for use throughout the federal government remain. Some agency attorneys have opined that the multitude of laws and rules pertaining to data sharing for different data sets preclude standard/templated MOUs. A framework and clear direction from OMB may address this constraint. Regardless, legal research and cultural changes are needed to develop a path forward.

## Expanding Secure Access to CIPSEA Data Assets

Title III of the Evidence Act requires OMB to issue a regulation that provides standards for (1) categorizing data sensitivity levels, (2) identifying a tiered access framework that enables agencies to improve data access, and (3) conducting comprehensive risk assessments. For more information, see Table 1. CIPSEA 2018 Regulations.

During its second year, the Committee engaged with a wide variety of experts, including from the ICSP and other entities inside and outside the federal government, on topics related to the forthcoming regulation on expanding secure access to CIPSEA data assets. This section summarizes the key takeaways from this investigative process and provides a series of recommendations on the regulation, touching on the content of the regulation, implementation strategies, and implications for governments at other levels, including the role for the NSDS. The Committee notes that, just like the pieces of the regulation, these recommendations are interrelated and should be considered—and implemented—as a set.

### *Regulation Content: Risk and Utility Framework*

**Recommendation 1.6. OMB should adopt a risk-utility framework as the basis for standards on sensitivity levels, access tiers, and risk evaluations as part of the regulation on expanding secure access to CIPSEA data assets.**

Risk and utility should be separately measured but jointly determined. Evidence on data use should be used to inform the measurement of value, models from the public and private sectors should be used to measure risk, and the Five Safes framework should be applied to develop combined risk-utility metrics that are open and transparent. Key considerations on these aspects are described below.

**Utility.** Because there is limited information on data use, agencies have historically relied on a variety of methods, including standing advisory committees of expert data users, data user surveys, and literature reviews of citations, to find out how their data have been used. In addition, the lack of an automated method to search for and discover what data sets are used in empirical research leads to fundamental reproducibility challenges, threatening the utility of these data for research. The Committee discussed examples of how evidence on data use can inform the measurement of value and, by extension, be used to increase value, including ACDEB’s use cases, the [Democratizing Data project](#) (also known as “rich context”), [measures of conservation value in the forestry service](#), and the [automation of research workflows](#) in Federal Statistical Research Data Centers and elsewhere. For more information, see Appendix B. ACDEB Use Cases and the Supplemental Information—Other Models and Examples [posted](#) with this report.

**Risk.** The Committee explored well-tested frameworks and tools to measure and mitigate risk and examined tools to assist organizations with risk assessments. Risk assessments provide a key method to evaluate the information a data set contains while weighing and evaluating the value and benefits against potential privacy risks associated with a release. For more information, see the Supplemental Information—Other Models and Examples [posted](#) with this report.

**Combining risk and utility.** The [Five Safes](#) framework provides an appropriate approach to operationalizing value and risk metrics. Box 3 provides high-level considerations of this framework.

### Box 3. The Fives Safes

The Five Safes framework takes a multi-dimensional approach to balancing disclosure risk and utility when using confidential data for evidence building and research. Each “safe” (safe projects, safe people, safe settings, safe data, and safe outputs) refers to an independent but related aspect of disclosure risk.

Such a framework offers a standardized approach for researchers, other data users, and consumers to create new products and data assets. This framework facilitates data exchange and offers advantages such as interoperability and modularity by design. Therefore, it has the potential to reduce the total cost of ownership, provide opportunities to reduce burden, and fuel new data sharing.

The Five Safes framework is broadly applicable. For example, the Committee discussed the five safes as a series of metrics to weigh disclosure risk and utility, explored each of the five dimensions as “guardrails” to guide decisionmaking around technical infrastructure, and investigated technical approaches to implement each safe. The state and local government representatives on the Committee agreed that services aligned with the Five Safes framework could encourage greater state participation. The five safes are an ideal framework for the NSDS.

For a general overview of the framework, see the Supplemental Information—Other Models and Examples [posted](#) with this report.

Statistical agencies can demonstrate value to data providers and users by enabling safe projects, people, and data.

- **Safe projects.** Statistical agencies can identify high-value projects that advance the goal of evidence-based policymaking by working closely with federal programmatic agencies and state and local government agencies. Utility measures can be identified by the agency, researchers, and stakeholders and validated through automated measures of data usage.
- **Safe people.** Agencies can increase the size of the skilled workforce available to produce evidence through standards for accrediting safe researchers in a transparent and accountable manner. Learning agendas can be expanded to include more hands-on training in privacy and confidentiality.
- **Safe data.** Agencies can increase access to safe data—both to safe confidential microdata and to safe summary data.

Statistical agencies can mitigate risks using restricted data by implementing the Five Safes framework, particularly involving safe people, settings, data, and outputs.

- **Safe people.** Agencies can institutionalize and operationalize the concept of safe people by creating open and transparent access policies that specify who can read or edit data, for what purposes, and how derived data products may be shared. Indemnification could be expanded to include analysts accessing confidential summary tabulations, as has been implemented by the Midwest Collaborative.

- **Safe settings.** Agencies can enable more modes of data access through secure remote environments and privacy-preserving technologies.
- **Safe data.** Agencies can offer tools and technical assistance to produce safe data, including hashing algorithms and guidance on producing synthetic data.
- **Safe outputs.** Agencies can support production of safe outputs, including confidential summary tabulations, synthetic data, traditional statistical disclosure limitation techniques, and evolving methods such as differential privacy.

### **Regulation Content and Implementation Strategy: Tiered Access in Practice**

During its second year, the Committee engaged with a wide variety of experts, including from the ICSP and other entities inside and outside the federal government, to understand (1) the need for better data access and (2) options to make data more accessible. OMB should use these insights to inform guidance on tiered access as part of the regulation on expanding secure access to CIPSEA data assets.

These experts expressed many needs and challenges related to better data access, including the following:

- **Needs.** Needs include the ability to access timely, relevant data with sufficient granularity to inform policy development and decisionmaking at all levels of government and in the private sector. This often requires bringing together data from multiple sources—federal, state, local, and private—for purposes ranging from monitoring the local or regional impacts of economic, health, or environmental conditions to highly technical analytical analyses, including those using artificial intelligence approaches.
- **Challenges.** Challenges have several dimensions, including human capital, technology, and legal. There is a great deal of competition for skilled data scientists, analysts, statisticians, and social scientists, and many in the current workforce do not have sufficient training to use modern analytical methods and tools. Access to technology and tools varies widely across levels of government and among other users. Technological challenges impact both data users and data providers. Lack of data standards and common metadata limit data interoperability. Data collected for administrative or statistical purposes may have restrictions on who, where, and for what purposes they can be used. Data owners face the challenge of responsibly balancing the need to protect the privacy of data subjects with the utility and accuracy of information made available to the public. For more information, see Risk and Utility Framework above.

There is a range of technical solutions that can help address some of these challenges. In particular, the use of data access tiers is a way of expanding access while implementing the [principle of least privilege](#), that is, granting access to only those data assets absolutely necessary to meet a user’s needs. Two extreme access tiers can be defined as restricted data (i.e., available to agency personnel, contractors, and approved agents) and open data (i.e., with no access controls, available to the public). In between these extremes are other methods that offer the promise of expanding access to data in ways that are timely and provide enhanced data protection. The Committee heard from experts on current and emerging access modes, some of which are described below.

**Restricted data tier.** The restricted data tier is designed to satisfy each dimension of the Five Safes framework. Currently, data enclaves provide authorized users direct access to curated microdata in a safe, secure setting managed by or for data owners. Traditionally, restricted data were accessed in specially outfitted data centers located in universities or federal agencies. Increasingly, these enclaves also offer virtual access, relying on technologies to monitor users. Consortia, led by nonprofit or academic institutions that leverage cloud technology, are bringing together a wide range of data producers and users to solve important social problems. Enclaves often provide access to state-of-the-art analytical tools, and some provide training for data users. The Committee heard about several examples including the following:

- [Inter-university Consortium of Political and Social Research \(ICPSR\)](#);
- [Administrative Data Research Facility](#) provided by the Coleridge Initiative;
- National Opinion Research Center ([NORC](#)) at the University of Chicago;
- [Pediatric Cancer Data Commons](#) at the University of Chicago;
- [Federal Statistical Research Data Center \(FSRDC\)](#) network; and
- Virtual data enclaves at the NCHS [Research Data Center](#) and at the Bureau of Labor Statistics—[National Longitudinal Surveys of Youth](#) and [Census of Fatal Occupational Injuries](#).

Licensing agreements can also provide access to restricted data but place the responsibility for ensuring a safe setting on the data users, who must follow conditions and protocols specified in the license agreement.

**Open data tier.** In contrast to the restricted data tier, which implements all dimensions of the Five Safes framework, the open data tier requires protocols and data processing, reducing the need for vetting and specifications involving safe people, settings, and output. Open data can be used by everyone. There is an inherent tradeoff between the quantity and quality of data released and the level of privacy protection afforded to the data. Open data processing may involve establishing minimum cell sizes for tabular releases, coarsening responses by collapsing categorical data, or adding noise to estimates. Agencies have traditionally relied on a combination of techniques to protect public data releases, adjusting the disclosure limitation strategy to fit perceived sensitivity of specific variables or even subsets of the population. The Federal Committee on Statistical Methodology’s [Data Protection Toolkit](#) provides resources to apply best practices in this evolving space.

Examples of open data include the following:

- Summary tabulations
- Public-use data sets
- Synthetic data

Due to advances in computer technology and analytical techniques, intruders are increasingly capable of revealing information about individual records by combining multiple data sources. This growing threat has required federal agencies to explicitly grapple with the tradeoff between privacy and utility in their public releases. Technical approaches such as differential privacy consider and quantify how much information about data subjects “leaks” out of each publication or release of statistics. These methods are currently implemented by the [Census Bureau](#) and by private sector organizations. While the science around privacy preservation is advancing quickly, the Committee heard concerns from data users about a lack of transparency on the risk and threat models, how the methods are implemented, and a lack of tools to understand the noise added, suggesting the need for additional research.

***Middle data tiers.*** Federal statistical agencies have a long history of producing open data products, mostly limited to information collected as part of each agency’s mission. Increasingly, users need data that combine information from multiple sources to form a more complete understanding of a subpopulation or economic segment. The Committee examined efforts such as the Census Bureau’s [LEHD](#) program, the work of [Opportunity Insights](#) at Harvard University, and the Coleridge Initiative’s [Unemployment to Reemployment Dashboard](#). These examples illustrate the promise of additional access tiers. Combining multiple data sets can provide insights not possible from any single data source in isolation.

There can be legal restrictions, however, on physically sharing data. The Committee heard from cybersecurity experts who noted that storing multiple data sets in a single data center creates an attractive target for hackers. Emerging privacy-preserving technologies, such as secure multiparty computation, which access data *in situ*, offer promise for expanding insights from data while minimizing disclosure risk.

“Middle” tiers provide users with indirect access to microdata and return outputs that have been processed to ensure protection of the underlying data. Examples include the following:

- ***Controlled access online data tools.*** These tools provide a computing environment that allows *approved* users to run tabulations or queries on restricted-use data through an online portal.
- ***Validation server.*** This is a type of open access query tool designed to validate research findings developed using synthetic (open) data. This combination can greatly increase the impact and usefulness of synthetic data.

## ***Implementation Strategy: Increase Adoption of Privacy-Preserving Technologies Within Government***

**Recommendation 1.7. OMB, in coordination with the ICSP, should promote the use of privacy-preserving technologies in the tiered access framework required under Title III of the Evidence Act by identifying an initial set of promising tools over the next 1 to 3 years.**

Through its own expertise and information gathering with outside experts, the Committee discussed how governments at all levels are seeking answers to questions based on input data they never “see.” The use of privacy-preserving technologies (PPTs) in new access tiers will enable evidence building while balancing input and output privacy needs.

There has been a rapid advance in maturity and pilot deployments of PPTs, as defined by the [United Nations Handbook on Privacy-Preserving Computation Techniques](#). These tools include secure multiparty computation, homomorphic encryption, zero knowledge proofs, and trusted hardware enclaves. Examples that the Committee explored include the following:

- Differential privacy has been deployed by [Apple](#), [Google](#), [Facebook](#), the [Department of Education](#), and the [Census Bureau](#).
- Several deployments and pilots, like the aggregations performed in the [Boston wage gap study](#) and the privacy-preserving record linkages conducted by the [Department of Education](#) and the [National Center for Health Statistics](#), and in [Allegheny County, Pennsylvania](#), provide evidence that secure multiparty computation is mature for use in some evidence-building applications.
- Homomorphic encryption and zero knowledge proofs have also been adopted within several recent and ongoing pilot projects, such as within federated learning algorithms for financial crime detection and pandemic forecasting in a prior [U.K. TechSprint](#) and an ongoing [U.S.-U.K. prize challenge competition](#).

For other examples of the adoption of PPTs within and beyond the statistical system, see Appendix D. OMB/ICSP Workstreams, Appendix E. ACDEB Subcommittee Guest Speakers, and the Supplemental Information [posted](#) with this report.

PPTs can be an alternative to physically locked-down sites; they require less movement of people and data and offer the potential for a more federated and connected world where data “lives” at the source agencies, thus increasing privacy. Continued research is needed to develop solutions that address more complex research questions and to understand the social, legal, and technical impacts and potential risks of using these technologies, as discussed in a [JASON report](#) and responses to an OMB Office of Science and Technology Policy/National Science and Technology Council Subcommittee on Networking and Information Technology Research and Development [request for information](#).

Depending on the application, these new technologies may require substantially more time and computational power than existing approaches. These costs and investments may limit PPT application and deployment at scale, which is why exploring and piloting these applications is necessary. It is crucial to assess these costs with alternative modes of access, such as FSRDC expansion or maintenance of validation programs for synthetic data. These technologies must have standardized and cleaned data up front, requiring comprehensive and accurate metadata from data providers.

Based on this evidence, which reflects a subset of current applications across the public and private sectors, the Committee asserts that some PPTs are mature to deploy at the scale for select statistical applications. These tools can be used readily to answer many questions—though not all—just like FSRDCs, other enclaves, and current synthetic data programs. Furthermore, no single solution will address all questions or needs, and there is a need for additional research and development. The Committee encourages research across PPTs to identify combinations of new and traditional approaches to provide safe data in safe settings. For information on the role of the NSDS to advance the development and deployment of PPTs, see Recommendations Part 3. NSDS Functions—Research and Development.

### ***Regulation Content and Implementation Strategy: Shared Responsibility and Safe Harbor Principles***

**Recommendation 1.8. OMB, in coordination with the ICSP, should identify models for shared responsibility among data providers and users and provide guidance on applying such models through the regulation on expanding secure access to CIPSEA data assets.**

The Committee affirms the importance of establishing standards to assess data assets in terms of sensitivity levels, corresponding levels of accessibility, and whether less sensitive versions of the data can be created, as required by Section 3582 of the Evidence Act. The Committee concludes that the regulation should reflect a set of guiding principles, including that (1) disclosure risk is on a continuum and is not binary, (2) not all data are equally sensitive, (3) there is shared responsibility between the statistical agency and users for protecting and not disclosing or re-identifying data, and (4) there is a need to protect good faith actors (i.e., data providers and users who take all precautions appropriate for known risks).

Agencies and stakeholders have typically operated in a binary legal environment where data are either accessible or not, which drives risk-averse behavior. The regulation on expanding secure access to CIPSEA data assets provides the opportunity to set policies that foster more open legal interpretations and facilitate the needed cultural change that recognizes disclosure risk as a continuum. This is an important step toward encouraging agencies to look for acceptable levels of risk. This would free agencies from interpreting the law as requiring risk to be zero, thus enhancing their ability to expand data access. This practice is shifting, as some agencies are using privacy-preserving approaches to expand data access.

This principle is reinforced by the idea of shared responsibility. If data users could share the responsibility not to re-identify data, that could further buffer against the need for zero risk. Mandating good faith behavior is a challenge. Additionally, to foster a culture of innovation, there needs to be protection for good faith actors, whose ethical hacking can alert organizations to vulnerabilities. For instance, regulations could provide safe harbor for “[good-faith security research](#)” as defined by the Library of Congress. It is important to shift the risk posture away from resting solely with statistical agencies and toward encouraging good faith actors to develop tools for expanding data access while protecting privacy and confidentiality. For more information, see Appendix D. OMB/ICSP Workstreams and the Supplemental Information [posted](#) with this report.

## ***Regulation Content and Implementation Strategy: Strengthen Penalties for Inappropriate Access and Use of Federal Data Assets***

### **Recommendation 1.9. OMB should identify new frameworks for determining sanctions for unauthorized access, use, and disclosure of government data assets.**

As part of these new frameworks, OMB should develop approaches that tie penalties to data sensitivity levels and access tiers, as outlined in the regulation on expanding access to CIPSEA data, and that consider different penalty structures for different actors—such as individuals and institutions.

While the Evidence Act outlines potential criminal and civil penalties for willful disclosure of restricted, confidential data assets, such traditional penalties applied in traditional ways may not be effective in protecting these data assets from unauthorized access and use by a broader set of actors. For example, these penalties focus on the disclosure of confidential microdata accessed in a secure enclave, which may be too limited to cover the tiered access approach envisioned under CIPSEA 2018. Furthermore, penalties are directed toward individual actors, not institutions who may have a large incentive and significant resources to leverage government data, in combination with a plethora of information available from other sources, potentially for financial gain.

The Department of Justice (not the statistical agencies) decides whether to prosecute a case; however, a more systematic penalty structure that focuses on new ways to disincentivize unauthorized access and use could (1) serve as an effective deterrent for inappropriate access and use and (2) support shared responsibility among statistical agencies and data users for protecting government data assets. For more information, see Shared Responsibility and Safe Harbor Principles above.

The Committee recognizes that new frameworks could include administrative penalties and other mechanisms for discouraging misuses or abuses of confidential data in this broader context. The Committee specifically recommends that this framework be further explored to align with the forthcoming public trust regulation required under the Evidence Act. For more information, see Responsibilities of Statistical Agencies and Trust above and the Supplemental Information—Other Models and Examples [posted](#) with this report.

## The Standard Application Process

### *Regulation Content and Implementation Strategy: SAP Evolution*

#### **Recommendation 1.10. OMB should prioritize resources to optimize the user experience for the Standard Application Process.**

The Committee applauds the efforts to date to launch a user-designed Standard Application Process (SAP). Developing the SAP is an important and necessary effort, mandated by Title III of the Evidence Act. The coordination required across the federal statistical system to launch such an effort is challenging. By law, the SAP must establish clear procedures around application submission and review, timelines, public reporting requirements, and an appeals process. All recognized statistical agencies and units are required to use the SAP to approve access to restricted-use data assets. OMB must routinely seek stakeholder feedback to ensure that the SAP is useful by identifying opportunities to adapt functionality as user groups, technologies, and resources come online in the years ahead.

During Year 2, the Committee discussed key elements of the [proposed SAP policy](#), including the following:

- The goal of the SAP is to reduce barriers and streamline the access process without increasing disclosure risks.
- The SAP must provide an identical application process, a common application form, criteria for granting access, timeframes for determinations, and an appeals process.
- The process includes approvals for safe people and safe projects.
- A single portal will not solve all problems; however, it is an important first step and an improvement over the current state, where users must rely on a distinct process for each data set when requesting data from multiple agencies.

The Committee also explored how the SAP and related activities could evolve. Enhancements for the approval processes, technical assistance, and integrating the SAP with the NSDS are described below.

***Approval processes.*** The SAP intersects with other CIPSEA regulations to create a standard framework that aspires to offer access to administrative data of nonstatistical federal agencies. The SAP should work in conjunction with the required data sensitivity levels and access tiers.

Users could be approved in a way that allows them to be involved in multiple projects without having to redo paperwork or that offers an expedited clearance process—like pre-clearing federal employees who may need access to sensitive data. To make this happen, OMB must standardize requirements for who can access data, as these differ by agency. Projects could also be approved so that they could be re-run at regular intervals, based on the same data sets and parameters, without re-applying each time.

The Committee acknowledges that friction in the approval process is helpful for protecting privacy and confidentiality, a desirable inefficiency. In addition, it may seem attractive to expedite clearances and project approvals for certain groups or types of research, but this could introduce inequities into the process. Shortening the cycle time for project approval while mitigating bias could be achieved by increasing dedicated resources for proposal review and data editing at the agency level. For more information, see Recommendations Part 6. Resources and Funding.

**Technical assistance.** As a start, the SAP will include agency points-of-contact so that users can inquire about specific data sets and uses; however, there is a larger need for data concierge services that connect users to the right data and assist with developing data requests. The NSDS could offer technical assistance like this, acting as a coordinator between data users and data providers. For more information, see Recommendations Part 3. NSDS Functions—Coordination (Support a High-Quality User Experience, Including Providing Technical Assistance).

**Integration with the NSDS.** The SAP should integrate with the NSDS so that users can seamlessly apply for access to diverse data assets—whether those data are nonpublic assets of statistical agencies, programmatic data acquired by statistical agencies from other federal agencies through the Presumption of Accessibility, or administrative data provided by state and local governments for their own or others’ evidence-building needs. Further, the Committee recognizes that the long-term benefits of a more integrated approval process would offer substantial gains for the evidence ecosystem and encourages OMB, the ICSP, and administrative programs to explore how to move in this direction more expeditiously.

The Committee acknowledges that there are challenges to establishing the SAP and integrating it with the envisioned NSDS, including the following:

- The SAP is currently focused on data discovery for a subset of federal statistical agencies. These agencies have prioritized their statistical data collections but not necessarily their administrative data holdings. Coverage in the SAP needs to expand to include federal statistical agencies and units, programmatic agencies, and state and local partners who are ready and willing to share data.
- While the proposed SAP policy standardizes elements of the approval process, the timelines for agency response would still be lengthy—12 weeks for access to data from one agency and 24 weeks for data from more than one agency. The process may simply be too long to provide timely and actionable evidence for decisionmaking, especially when policymakers are grappling with crises like the COVID-19 pandemic.

- The clearance of individuals is currently conducted outside of the SAP because different agencies have different criteria. In the short-term, there are clear necessities for federal agencies to maintain strict control of final approvals, but it would be helpful to see a process and dialogue for federal agencies that examines the sensitivities, nuances, legal restrictions, and financial costs of obtaining clearances as a means for accelerating future approvals in certain cases. In addition, the process must be flexible enough to meet all applicable legal and regulatory requirements for states to participate. For example, currently, states complete their own background checks using preferred vendors. To encourage states to participate, the SAP must implement a clear and acceptable standard for meeting security needs and must establish best practices that reflect a level of confidence.
- The use of the SAP for non-CIPSEA entities to provide access to their data assets would be voluntary, except as required under the Presumption of Accessibility. Therefore, for additional federal agencies and governments at other levels to embrace an expanded SAP, the ICSP must demonstrate the effectiveness of the process—both for increasing data access to provide value and for protecting data assets to mitigate risk.
- Meeting required SAP timelines could be a challenge for data providers at federal programmatic agencies and state, territorial, local, and tribal governments who may not have established procedures or dedicated staff for reviewing applications.

It is critical that key stakeholders from the public and private sectors have a voice in the buildout of the SAP, both as data users in the initial phases and, potentially, as data providers. In addition, OMB and the ICSP should assess the risk of expanding data access versus the utility of increased access and should collect metrics on application denials, the number and percent of proposals approved and denied, and cycle time for approvals to guide improvements. For transparency, these metrics should be available to the public. For more information on utility and risk metrics, see Expanding Secure Access to CIPSEA Data Assets (Risk and Utility Framework) above and Recommendations Part 4. NSDS Organizational Structure and Governance—Holistic Governance Approach to Transparency and Accountability.

## Recommendations Part 2. Other Evidence Act Items

### Sharing of Business Data Among Designated Statistical Agencies

**Recommendation 2.1. OMB should propose legislative text as part of the FY 2024 Budget to allow full implementation of Evidence Act Section 3575 for the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the Census Bureau.**

CIPSEA 2002 includes a provision that “authorizes the sharing of business data among the Bureau of the Census, the Bureau of Economic Analysis, and the Bureau of Labor Statistics for exclusively statistical purposes.” The Committee acknowledges that despite this provision, the goals of such “data synchronization” have not been realized due to inconsistencies in [Title 13](#) and [Title 26](#). This provision is designed, among other objectives, “to improve the comparability and accuracy of Federal economic statistics by allowing [the agencies] to update sample frames, develop consistent classifications of establishments and companies into industries, improve coverage, and reconcile significant differences in the data produced by the three agencies.”

The Evidence Commission also recognized this longstanding issue: “Census is authorized to access Federal Tax Information for individuals and businesses under Title 26 (IRS statute). CIPSEA authorizes the Census Bureau, the U.S. Bureau of Labor Statistics, and the U.S. Bureau of Economic Analysis to share business data for exclusively statistical purposes. While these authorities permit the Census Bureau to access many different data sets, in practice data sharing agreements are negotiated separately and governed by applicable laws and policies.” The Committee acknowledges that CIPSEA 2018 not only reauthorized the “sharing of business data among designated statistical agencies” but also expanded the roles and responsibilities for statistical agencies to provide secure access to CIPSEA data assets, enabled by strong protections for privacy and confidentiality.

In 2022, the Biden Administration proposed to address this issue as part of the FY 2023 Budget. The latest proposal in the [Green Book](#) from the Treasury Department “would allow officers and employees of each of BLS, BEA, and the Census Bureau to access the same federal tax information (FTI) for businesses, and would permit BLS, BEA, and the Census Bureau to share such FTI amongst themselves (subject to restrictions).”

Congress intended to address the legislative barriers to sharing this information on multiple occasions. Small technical corrections in federal law, proposed by OMB, could finally enable important statistical activities to occur within the strong privacy framework of CIPSEA as envisioned more than 20 years ago.

## Federal Statistical System Learning Agenda

**Recommendation 2.2. The Federal Committee on Statistical Methodology should set a learning agenda promoting research on timely, relevant, and actionable questions, demonstrating the value and promise of the NSDS.**

Learning agendas, or evidence-building plans, have the power to “identify and address priority questions relevant to programs, policies, and regulations and to identify, prioritize, and establish strategies to develop evidence to answer important short- and long-term questions,” as described on [evaluation.gov](https://www.evaluation.gov). Furthermore, “learning agendas should be iterative, flexible, and transparent, and tailored to both meet an individual agency’s needs and to address agency-specific challenges to developing evidence.”

The Evidence Act casts a collective vision for the federal statistical system beyond the individual missions of each agency. The “[Leveraging Federal Statistics to Strengthen Evidence-Based Decision-Making](#)” Analytical Perspectives chapter in the President’s FY 2023 Budget proposal asserts that: “While each of the Federal statistical agencies, units, and programs have found innovative ways to address challenges individually, this individual approach is proving more and more difficult, and a successful future for the whole Federal statistical system will rely on increased collaboration.” The chapter also lays out a “blueprint for the future” that “relies heavily on the statutory framework provided by CIPSEA 2018 and operating as a seamless system.”

A learning agenda spanning the federal statistical system could be one mechanism through which the federal statistical agencies could advance a collective mission by harmonizing efforts around key issues and supporting accountability and transparency on these issues. To advance the vision for a seamless federal statistical system, NSDS staff or contractors could execute the system-wide learning agenda, and the ICSP could designate a lead agency for each pilot project.

The Committee recommends that a learning agenda be formulated to outline key learning objectives and priorities for a multi-year period that could guide system-wide needs. This could involve matters related to risk, privacy-preserving technologies, linkage methodologies, collaborations with the evaluation community, and engagement practices with state and local governments. The Committee recommends a participatory process for developing this agenda that involves key statistical system stakeholders in identifying core questions, objectives, and needs. The Federal Committee on Statistical Methodology may also look to partner with the Chief Data Officer Council and the Evaluation Officer Council for advice.

For more information, see Recommendations Part 3. NSDS Functions—Coordination (NSDS Project Sponsorship) and Recommendations Part 4. NSDS Organizational Structure and Governance—Near-Term Project Functions and Pilots.

## Recommendations Part 3. NSDS Functions

The vision for the NSDS—and its functions—has evolved over time. First, the Evidence Commission laid the groundwork for the NSDS and the services it would provide. The Evidence Act then shifted the landscape by assigning some of those functions to statistical agencies and requiring common frameworks without introducing a new entity; those functions are still evolving as the CIPSEA regulations and guidance are not yet completed. In that context, the Committee hones the role of the NSDS around four functions—(1) coordination, (2) communication, (3) research and development (R&D), and (4) data standardization. Looking to the future, the NSDS must continue to evolve with service demands, the emerging needs and capabilities of data providers and users, and the policy landscape.

The CIPSEA legal framework sets the baselines for how the NSDS must operate. In keeping with the codified responsibilities of statistical agencies, access to data through the NSDS will be for statistical purposes only. Data access, linkage, and analysis supported by the NSDS will never be used directly to enforce any local, state, or federal statute or to determine individual eligibility for benefits. Likewise, the NSDS must implement the standards for data sensitivity levels, tiered access, and risk assessments, as required to expand secure access to CIPSEA data assets. The NSDS must use the Standard Application Process (SAP) to receive applications for access to restricted-use data for evidence-building purposes.

This section presents recommendations related to the data service’s required capabilities. These functions are designed to advance evidence-based decisionmaking for diverse stakeholders throughout the data ecosystem.

### Function 1. Coordination

#### *Advance Evidence Act Implementation for the Federal Government and Beyond*

**Recommendation 3.1. The NSDS should coordinate with the ICSP, communities of practice, and other key stakeholder groups to establish best practices for implementing the requirements of the Evidence Act and should serve as a model for testing and demonstrating those best practices for governments at all levels.**

The Committee sees broad potential for the NSDS to promote best practices related to the Evidence Act, expanding on the Evidence Commission’s assertion that “the NSDS should be a center for excellence in developing, using, and sharing best practices for transparently and securely using confidential data for evidence building.”

Table 2 provides examples of potential NSDS activities that would serve as models for implementing the Evidence Act. For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance and the related recommendations and findings throughout this report.

**Table 2. Proposed NSDS Services to Support Evidence Act Best Practices: Examples**

NSDS Activities	Evidence Act Title and Section
Promote metadata standards, provide technical tools that harmonize the look and feel of data inventories and catalogs, and model the implementation of those standards and tools.	Title II (Sec. 3511) and Title III (Sec. 3582 and Sec. 3583)
Catalyze innovation in developing risk-utility metrics, data sensitivity levels, tiered access, and risk assessments.	Title III—Expanding Secure Access to CIPSEA Data Assets (Sec. 3582)
Promote the adoption of cutting-edge methods and privacy-preserving technologies.	
Provide tools and guidance that support data linkage.	
Develop file formats and standards that facilitate data sharing and analysis.	
Model data request review procedures that have integrity and are timely.	Title III—Standard Application Process (Sec. 3583)
Transparently report on the application process, project status, and linking and analysis results.	

## ***Support a High-Quality User Experience, Including Providing Technical Assistance***

### **NSDS Website as a “Front Door”**

**Recommendation 3.2. The NSDS website should serve as a “front door” to the nation’s data assets, organized around a set of personas that reflect basic user needs.**

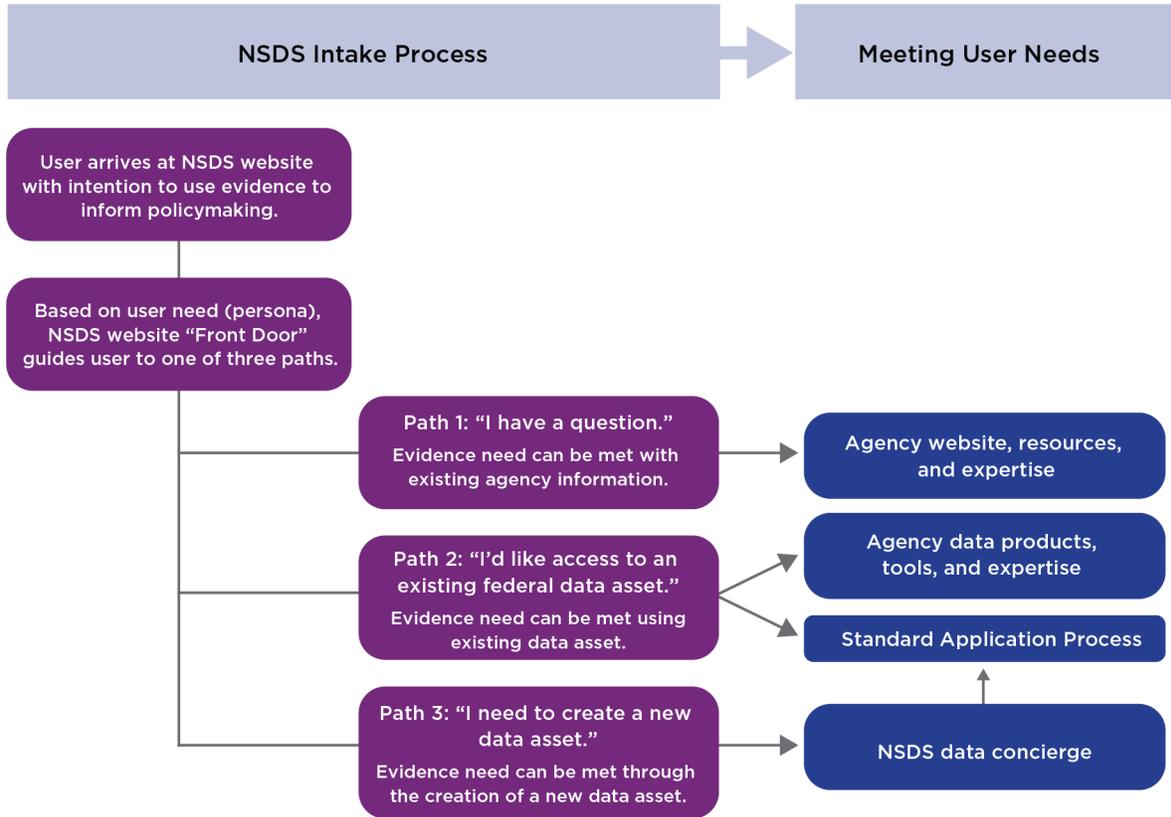
The NSDS website’s “intake process” and subsequent user experience should be organized around users’ needs. Furthermore, the user experience should be designed to integrate with the SAP and the Federal Data Catalog required under the Evidence Act as their functionality evolves. The NSDS website should supplement, not supplant, the many ways individuals discover and access data assets such as through search engines or existing relationships with statistical agencies. For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Standard Application Process and Recommendations Part 5. NSDS Technical Infrastructure and Tools.

Based on its fact-finding activities and the use cases identified by ACDEB members, the Committee believes that a potentially large number of individual users can be represented by a relatively parsimonious set of generalized personas. Although they may have other distinguishing characteristics, each persona may be thought of as having one of the following three distinct data access needs:

- ***“I have a question.”*** Some users will come to the NSDS for the purpose of locating one or more estimates (e.g., “What’s the median wage for a graduate of Great State University?”), unaware of extant tabulations that could meet their needs. The design of the NSDS should include the ability for users to identify a question domain (and subdomains as needed) and then be directed to existing agency resources that are likely to meet their information needs. If these resources fail to meet the users’ needs, the design of the NSDS should include the ability for users to contact (via a webform) agency subject matter experts who can provide further assistance.
- ***“I’d like to access an existing federal data set.”*** Some users will come to the NSDS with awareness of existing agency data resources, whether publicly available or restricted access. The design of the NSDS should include the ability for a user to identify a question domain (and subdomain as needed) or a specific agency and then be directed to a list of relevant data assets. Upon selection of a data asset, the user would be pointed to existing public-facing data products and analysis engines or, if needed, the SAP. This is consistent with Title II of the Evidence Act, which requires agencies to indicate in which access tier the data set is available and then direct users to the SAP if the data fall in a restricted tier. If this process fails to meet the user’s needs, the design of the NSDS should include the ability for the user to seek further assistance (via a webform) with existing tools from agency subject matter experts or with the SAP.
- ***“I need to create a new data asset by linking multiple data sets.”*** Finally, some users will come to the NSDS with the intention of linking multiple data sets and conducting original research, irrespective of the source of those data (e.g., federal, state, local, or proprietary) or their character (e.g., collected for a statistical purpose, generated from an existing business process). The design of the NSDS should route those individuals to a data concierge for assistance. For more information, see NSDS Data Concierges and Service Providers and Partners below.

Figure 1 describes the NSDS intake process and how it is designed to meet users’ needs.

**Figure 1. NSDS Intake Process and User Needs**



The ability of the NSDS to serve effectively as a “front door” to the nation’s data assets depends heavily on the resources and capacity of other actors and infrastructures in the evidence-building ecosystem. For example, if data concierges are to focus on assisting users with creating new data assets, then federal statistical agencies must be well-resourced to handle user demands for existing information, data products, and tools. The NSDS is not designed to meet every need for every user but to supplement other evidence-building activities and to coordinate efforts across the ecosystem. For more information, see Recommendations Part 6. Resources and Funding.

## Automation in the NSDS Intake Process

**Recommendation 3.3. The NSDS should identify opportunities for automation of its “intake process,” providing a high-quality user experience while focusing staff effort on complex user needs.**

Successfully positioning the NSDS as a “front door” to the nation’s data assets could result in potential users placing significant demands on the NSDS (or supporting agencies’) staff. In addition to optimizing the user experience of the NSDS site such that it is as self-service as possible, site designers should apply the latest automation technologies to meet simple user needs via virtual assistants.

Users of commercial websites often encounter “chatbots” or other forms of virtual assistants. While these assistants are typically capable of resolving only simple concerns, a well-designed chatbot may direct users to existing resources and perform common tasks, allowing human staff to focus their attention on more complex requests. Increasingly, these technologies are being used by websites maintained by U.S. government agencies. A notable example of this form of automation includes the Department of Education’s Aidan®, a virtual assistant used by Federal Student Aid at [studentaid.gov](https://studentaid.gov) to respond to questions from students and families about the Department’s federal financial aid programs.

### *NSDS User Experience and Technical Assistance Roles*

The Committee conducted a series of fact-finding activities designed to better understand the range of technical assistance and other support users might require. This included exploring offices within federal agencies tasked with supporting only those agencies’ employees (e.g., the National Institutes of Health Library Data Services), federally sponsored activities that make government data securely available to qualified users (e.g., the Federal Statistical Research Data Centers network and the Criminal Justice Administrative Records System), and projects managed by academic and research organizations (e.g., the Inter-university Consortium for Political and Social Research, the Coleridge Initiative, and the National Opinion Research Center).

### *NSDS Data Concierges and Service Providers and Partners*

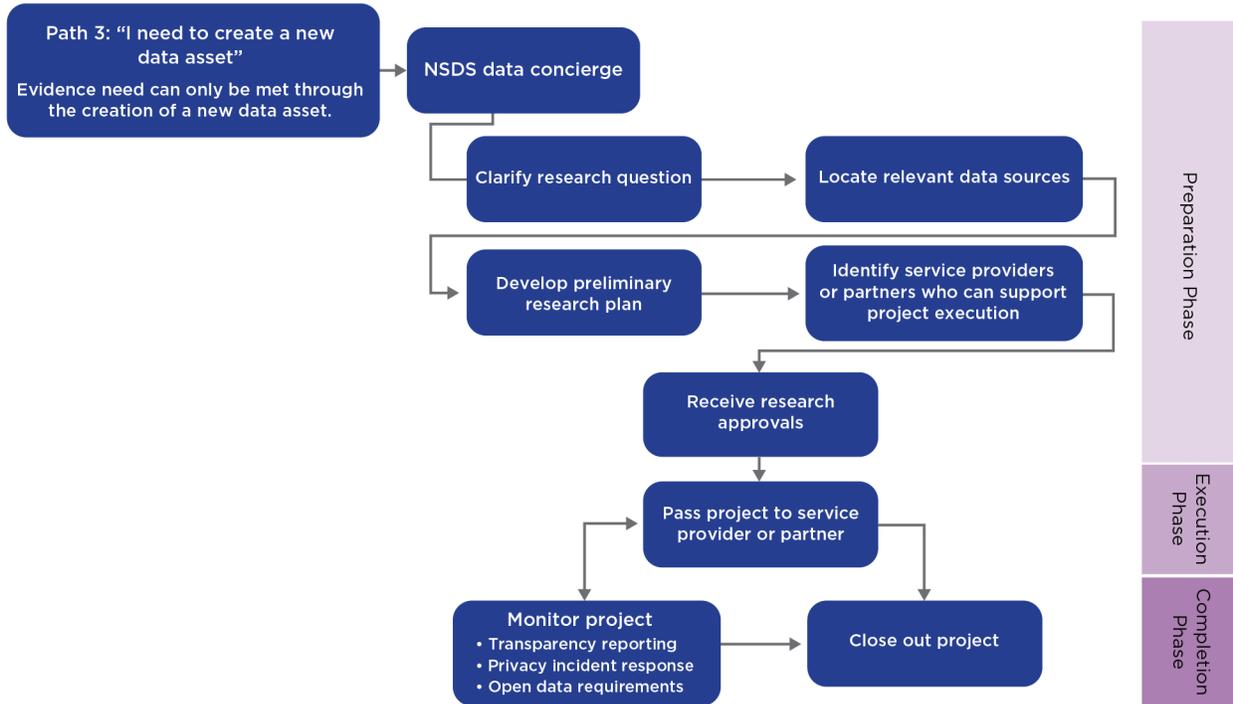
The Committee conceptualizes the experience of NSDS users who need to create new data assets (i.e., path 3 described above) as consisting of three phases: (1) a preparation phase, (2) an execution phase, and (3) a completion phase.

To support the preparation and completion phases, the Committee recommends the following:

**Recommendation 3.4. The NSDS should employ data concierges to help users refine their research projects, discover relevant data, and acquire access to that data.**

The execution phase is the province of what the Committee refers to as “service providers and partners,” that is, entities within the evidence-building ecosystem that provide analytical and other resources needed to complete user projects. Figure 2 describes the NSDS technical assistance phases.

**Figure 2. Technical Assistance Phases**



**Phase 1 (preparation phase) data concierge activities.** The preparation phase begins with clarifying a user’s research question and concludes with the user’s submission of a completed research proposal to the appropriate authority for approval. Typical activities, and the resources needed to complete them, include the following:

- Clarifying the initial research question, including referring the user as appropriate to existing and available data assets that could meet the user’s needs, whether those data are from federal statistical or program agencies, from governments at other levels, or from the private sector. This requires data concierges to have knowledge of the data sets, access to the metadata, and an understanding of the physical and computational security requirements of the data sets. In addition, data concierges must have connections with subject matter experts to identify potential sources of data, their fitness for purpose, and relevant access procedures. While the access process varies for different data sets, the goal is for the SAP to integrate with the NSDS and create a seamless experience for users applying to access any data asset for an NSDS project. For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Standard Application Process.
- Supporting the development of all aspects of NSDS research proposals, including (1) assistance in preparing relevant access agreements prior to their submission to the appropriate authority for approval and (2) aiding users in obtaining access for approved uses. This requires that data concierges be aware of analytical resources within the NSDS and their suitability for executing the user’s research proposal. Data concierges must also have connections to agency subject matter experts responsible for approving the use of data for statistical purposes.

- Handing off to one or more service providers or partners who will assist the user in the execution of the project. This requires data concierges to have an effective working relationship with NSDS service providers and partners.

***Phase 2 (execution phase) responsibilities of service providers and partners in the data ecosystem.*** Just like the mission and work of the federal statistical system, the remit of the NSDS data concierges is broad and complex. As such, the Committee recommends that certain services to support approved NSDS projects be contracted with service providers and coordinated with partners, including other statistical agencies. Those services include the following:

- Providing computing and network environments, including associated cyber and physical security responsibilities;
- Tiered access services, including aiding data contributors in deciding which access controls to apply to their data sets;
- Data cleaning services and metadata services, including metadata standardization efforts;
- Analysis and reporting services; and
- Disclosure avoidance and review services.

***Phase 3 (completion phase) data concierge activities.*** This phase includes ongoing monitoring of users' projects as they are executed by NSDS service providers and partners and any relevant project closeout activities. In addition to monitoring project status (e.g., adherence to proposed project milestones and deliverables), activities in this phase include the following:

- Fulfilling transparency requirements, such as regularly updating public-facing project registries containing information like (1) project names, descriptions, and anticipated value to the public; (2) project teams and affiliations; (3) data used; (4) research questions and methods; and (5) anticipated deliverables and associated delivery dates, in keeping with the requirements established by the SAP policy. This information should be made publicly available. For more information, see Recommendations Part 5. NSDS Technical Infrastructure and Tools—Searchable Project Inventory.
- Documenting the actions of researchers, service providers, and partners taken to meet requirements for various access tiers under the regulation on expanding secure access to CIPSEA data assets.
- Promoting principles of open and reproducible science, as appropriate (e.g., the [FAIR principles](#)), including (1) assisting researchers with documenting research so that it may be reproduced, including developing code and output to satisfy peer review standards, and (2) describing the origins of the data sets and procedures used in data analysis, upon request. For more information, see Recommendations Part 5. NSDS Technical Infrastructure and Tools—Auditability.

As with the automation of the intake process described above, the NSDS should explore mechanisms to automate data concierge activities (such as using bots) when feasible. In addition, the NSDS should publish usage metrics for both automated and direct assistance to gauge service demands and unmet needs.

## Technical Assistance Leads

**Recommendation 3.5. The NSDS should employ technical assistance leads who develop educational resources for data providers and data users related to the methods and technologies used by the NSDS and in the broader evidence-building ecosystem.**

As a result of the fact-finding activities described above, the Committee recommends that NSDS technical assistance leads make a variety of education and training opportunities available to the user community. Activities should be offered in a variety of formats, including (1) one-on-one consultation services, (2) synchronous learning opportunities, and (3) asynchronous learning objects.

Education and training topics should reflect needs identified through systematic engagement with federal, state, territorial, local, and tribal partners and their collaborators. For more information, see Systematic Engagement for Continuous Improvement below.

Relevant topics could include, but are not limited to, the following:

- Basic NSDS operational protocols (e.g., how to write and submit a research proposal);
- Data and metadata quality, management, and cleaning;
- Data linking methodologies;
- Privacy-preserving techniques;
- Legal issues associated with data sharing and linkage; and
- Disclosure avoidance.

Box 5. Complementary Training for Lawyers and Data Staff provides one example of the type of targeted training the NSDS could offer.

## Box 5. Complementary Training for Lawyers and Data Staff

During its second year, the Committee heard many examples of how legal interpretations tend to push agencies to prevent, rather than permit, expanded access to data assets, despite the Evidence Act’s mandate to the contrary.

The Committee also recognizes that challenges like these could be overcome, in part, by engaging with legal staff when developing new programs or data-sharing agreements and by educating both lawyers and program experts on the intent and possible applications of the law. The National Center for Science and Engineering Statistics’s use of its “Other Arrangements” Authority to stand up America’s DataHub Consortium (ADC) is one example of how partnering with legal counsel to surface new ideas, explore options, and make data sharing and programmatic decisions can yield many benefits that impact the broader evidence ecosystem. The application of this authority opened new and flexible contracting opportunities that the ADC may not have been able to realize without the support of the Office of General Counsel. Actionable Intelligence for Social Policy at the University of Pennsylvania has developed guidance to inform lawyers reviewing data linkage projects.

As part of its comprehensive communication and education strategy, the NSDS could provide educational resources and support training on data sharing for agency lawyers at the federal, state, and local levels. For maximum effect, this training should be conducted by lawyers and should clearly describe the importance of tools like common templates and approaches. Likewise, the NSDS could provide resources and training for data staff on legal issues around data sharing. Offering this type of training would help to build a common understanding of the legal challenges and possibilities of data sharing within and across levels of the government.

For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Presumption of Accessibility for Statistical Agencies and Units (MOUs) and Recommendations Part 4. NSDS Organizational Structure and Governance—America’s DataHub Consortium as the Foundation for the NSDS.

The Committee recommends that the technical assistance leads, with the support of the NSDS data concierges, also develop resources that support users in the design and execution of common analyses, such as matching education records to workforce or social benefit records. Resources could include research proposal templates and illustrative code for common data management, analysis, visualization, and reporting tasks.

## Crowdsourced User Support

**Recommendation 3.6. The NSDS should actively encourage the development of communities to crowdsource support for users, complementing the work of the data concierges.**

The NSDS should support the growth of user communities related to the data service’s operations. These communities might already exist, be expanded, or newly arise around shared disciplinary interests (e.g., postsecondary education and training), common methods (e.g., spatial analysis), or other technical interests (e.g., secure multiparty computation). Communities might take the form of relatively traditional structures with familiar incentives for participation, such as role-alike communities of practice, or less traditional forms, such as code competitions (e.g., the Coleridge Initiative’s [Show US the Data](#) competition).

Potential activities include the following:

- Shared code for common data management, analysis, visualization, and reporting that complements material created by NSDS staff;
- A repository of working papers that illustrate the use and value of NSDS-based analyses; and
- Discussion forums that create a space for members to offer assistance to others in the user community.

## Project Sponsorship

**Recommendation 3.7. The NSDS should regularly sponsor projects that demonstrate the value of streamlining data sharing and increasing coordination, specifically with projects that highlight cross-functional, cross-agency, and cross-governmental topics.**

In addition to assisting others with developing and carrying out project ideas, as part of its broader coordination function, the NSDS should proactively coordinate evidence-building efforts across federal, state, and local governments, as well as with the private sector. By sponsoring projects, the NSDS could encourage stakeholders to partner on initiatives of shared interest with the aim of fostering better, broader, timelier, more efficient, and more collaborative outcomes.

The COVID-19 pandemic provides a unique moment for civic engagement and highlights the need for more comprehensive and cohesive development of data infrastructures and analyses not only at the federal level but at the state and local levels. ACDEB’s use cases and meetings with outside experts provide myriad examples of how the pandemic spurred more timely and relevant data for decisionmaking and spotlighted the need for enhanced coordination and systematic investment. For more information, see ACDEB’s use cases on education and workforce, health, and labor market activity and the summaries of outside meetings on the COVID-19 Research Database; the [National COVID Cohort Collaborative](#), and the Minnesota Labor Market Indicator Data Equity Pilots with the Bureau of Labor Statistics in the Supplemental Information [posted](#) with this report.

While progress has been made, federal efforts with states to improve their data infrastructures and analysis capabilities—like through the National Center for Health Statistics (NCHS) National Vital Statistics System modernization initiative, Department of Labor Data Equity Partnerships, or state-wide longitudinal data systems for education—are segregated across federal agency programs and, as such, are focused on specific domains, topics, or areas and do not tend to encourage collaborative efforts across the evidence ecosystem more broadly.

In addition to the near-term pilot projects sponsored as part of establishing the organizational structure, governance, and core capabilities of the NSDS, the Committee recommends that the NSDS regularly sponsor projects to incentivize collaboration within and across levels of government. These projects should encourage systems interoperability and data standardization and promote the use of standard formats for citing work. For more information, see Data Standardization below and Recommendations Part 4. NSDS Organizational Structure and Governance—Near-Term Project Functions and Pilots.

Furthermore, to sponsor projects, the NSDS must have a direct appropriation to provide sufficient resources as a sponsor across multiple program areas. For more information on the funding needs of the NSDS to support its mission, see Recommendations Part 6. Resources and Funding.

The Committee also discussed methods for ensuring that sponsored projects fulfill their approved objectives. One model, used by the Department of Education and other federal agencies, includes regularly reviewing grantee progress toward agreed-upon objectives and making the release of future years' funding contingent upon the achievement of key milestones. As part of establishing its project sponsorship program, the NSDS should explore mechanisms for ensuring compliance or performance, including funding structures.

The NSDS could also provide information that spotlights other potential project sponsors. For example, if a state official wants to link federal data to the state's own administrative records to evaluate a program, resources to complete the project could come from the state itself, a federal partner with a related interest in the project, or a foundation looking to advance the work. Similarly, the NSDS could assist state, territorial, tribal, and local governments in identifying private or non-federal entities that could provide resources to support data collection and infrastructure modernization efforts.

## Function 2. Communication

### *NSDS Mission Statement*

**Recommendation 3.8. OMB should adopt a clear statement of purpose for the NSDS that is rooted in its core value.**

The Committee concurs with the judgment of field experts that OMB must take the initiative to communicate clearly about the purpose of the NSDS. The Committee found that the initial Year 1 framing of the NSDS as a “philosophy, place, and service” resonated with those who were familiar with the effort, and broadly applies to the expanded collective mission of the statistical agencies under the Evidence Act. However, the Committee believes the “service” element—that is, the value the NSDS can provide and the support it can offer users at the federal, state, and local levels—is unambiguously the most powerful when expressing the potential of this distinctly new entity.

Although many possible purpose statements exist, the Committee suggests the following:

**“The National Secure Data Service benefits the American public by providing federal, state, and local partners the services they need to securely share, link, and analyze data to support the use of evidence in policymaking.”**

### *NSDS Public Information Hub*

**Recommendation 3.9. Through the NSDS website, NCSES should create a public information hub that serves as a central repository for resources about, and communications from, the NSDS.**

Fact-finding activities centered both on the exploration of potential governance structures for the NSDS and on approaches to communicating about the nascent data service suggest the importance of having a clear and comprehensive web presence for the effort. As such, the Committee recommends that the National Center for Science and Engineering Statistics (NCSES) establish an easy-to-find website for the NSDS that can serve as an authoritative source of information about the data service.

The purpose of the NSDS should be prominently featured, as should content related to the following:

- Potential (and, over time, actual) use cases that demonstrate both the data service’s value and how it operates, including spotlighting the risk-utility tradeoff and how determinations on data sensitivity levels, access tiers, and risk assessments are made. For more information, see Box 6. Communicate the Risks and Rewards of Data Use;
- The role of federal, state, local, and nongovernmental partners in the data service’s operations;
- Ongoing opportunities to learn more about the NSDS itself and, over time, about technologies, methods, and processes relevant to the operations of the data service;
- Frequently asked questions and other communications tools;

- An intelligent “front door” to the nation’s data assets that allows potential users to discover which data sets (including metadata) and features (e.g., routing to relevant published tabulations, access to publicly facing analysis tools to relevant data sets, beginning with the SAP) are most likely to meet their needs and how to access them. For more information see Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance (NSDS Website as a “Front Door”) above; and
- How to receive assistance from the NSDS data concierges. For more information, see Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance (NSDS Data Concierges and Service Providers and Partners) above.

### Box 6. Communicate the Risks and Rewards of Data Use

Addressing public perceptions of the privacy-utility tradeoff has become more difficult as the availability of new types of data, linkage technology, and the value of information has increased. This challenge is two-fold—the public does not fully understand the utility and relevance of government data, and government data are not used fully to create benefits for the public.

Education and transparency are the keys to increasing the public’s understanding of privacy risks and the appropriate tradeoffs between privacy and utility. That is, efforts to make data safer also make the data less useful. On the one hand, users could access the raw, confidential data with no disclosure restrictions, providing maximum utility but no privacy protections. On the other hand, no data could be disclosed, maximizing privacy but limiting utility. Users need to understand the tradeoffs between these extremes and how decisions are made to balance these tradeoffs.

On utility, access to data through the NSDS should start with a clear understanding of the value proposition for all parties involved by addressing key questions like: Why are stakeholders providing and combining data? And how could these efforts improve how agencies function and serve the public? The answers to these questions should be clearly communicated to the public.

In addition, historically, agencies have not clearly communicated about their methods to protect privacy and the quality of published data sets using traditional disclosure avoidance techniques. It is important to be transparent about these traditional methods, as well as newer frameworks built on the concept of differential privacy. This way, the public can assess the pros and cons of the different approaches. There also needs to be a mechanism that teaches a broad audience a whole new concept around privacy; however, there are no effective incentives to do this. One approach could be to build expectations regarding communications into grant programs.

The NSDS should clearly communicate with the public on the privacy-utility tradeoff and how determinations on data sensitivity levels, access tiers, and risk assessments are made in line with the forthcoming regulation on expanding secure access to CIPSEA data assets. The NSDS should also explore the potential to design surveys that capture public perceptions of the tradeoffs and hone communication strategies to better inform the public on these complex issues.

As part of its work to develop the NSDS web presence, OMB should engage NSDS stakeholders, including statistical agencies, to develop web standards. These standards should foster comparability between the NSDS and websites of data stewards and providers to ensure a nearly seamless handoff between the sites. These standards could eventually be adopted by statistical agencies and their partners. Doing so would promote a harmonized experience whether users are on the NSDS site or are directed to a site where a requested service or product can be accessed. The goal would be to reduce differences in look, feel, and functionality as much as possible.

### ***Comprehensive Communications Strategy***

#### **Recommendation 3.10. NCSES, in coordination with OMB and the ICSP, should build a comprehensive communications strategy for the NSDS.**

The experience of national organizations that advocate for the use of data and evidence in policy-making suggests that building awareness of, and support for, the work of the NSDS will require a comprehensive, sustained, and professionally managed communications strategy. The data service’s potential scope of work, the range of benefits it stands to provide, the sheer number of stakeholders with whom it could engage, and the extent of resources it will need to achieve its purpose are each significant. Investments in effective communication about the NSDS should be commensurate.

The Committee’s fact-finding indicates this strategy should include the following elements:

- The intentional use of the NSDS website as a focal point for others to learn about the data service’s purpose and benefits. For more information, see NSDS Public Information Hub above;
- Continued efforts to enumerate the groups (key audiences) with whom the NSDS should communicate;
- The identification of influential organizations that, by virtue of their connection to key audiences and other critical stakeholders, can efficiently and effectively amplify NSDS communications;
- The recruitment of champions from diverse communities, including all levels of government, who have the credibility and capacity to publicly advocate for the NSDS and its value;
- The development of communications artifacts (e.g., explainer documents, short use cases, talking points) and accompanying tools (e.g., social media posts, customizable presentation decks, prototype agendas for stakeholder meetings) that support efforts to educate the public and other key stakeholders about the NSDS and its benefits; and
- Communications activities that align with and span the phases of NSDS design and implementation, including the following:
  - ***Early-phase activities.*** These activities (e.g., virtual and in-person engagement events) precede an operational NSDS and focus on orienting stakeholders to the data service’s purpose, benefits, and broad design, inviting opportunities for dialog, feedback, and co-creation (“soft launch”);

- **Mid-phase activities.** These activities coincide with the initial launch of an operational and viable NSDS to generate enthusiasm and engagement, with the goal of yielding “quick wins” that, when documented and publicized, can accelerate NSDS use (e.g., proof points); and
- **Maintenance activities.** These are activities related to steady-state operations of the data service, highlighting a diverse array of projects (e.g., in-depth case studies on successful projects and those that fail) and inviting authentic and sustained engagement with stakeholder groups and user communities for the purpose of improving the data service’s operations. For more information, see Systematic Engagement for Continuous Improvement below.

### Systematic Engagement for Continuous Improvement

**Recommendation 3.11. The NSDS should build a system for routinely engaging with key partner groups and user communities for the purposes of needs sensing, operational improvement, and advocacy for the use of data to improve decisionmaking.**

For the NSDS to be useful and used, the data service must routinely, systematically, proactively, and authentically gather feedback from key partners for the purpose of improving its operations, including enhancing technical assistance services provided by NSDS data concierges and technical assistance leads. This is in keeping with the Evidence Act, including statistical agencies’ responsibility to maintain public trust and transparency requirements under the Standard Application Process. Audiences should include but are not limited to federal, state, and local data providers and users; researchers and research organizations; policymakers; and advocacy organizations.

Topics for engagement include the following:

- Evolving needs of data providers, including ways in which the data service could increase the value of the NSDS to federal, state, and local governments;
- Users’ experiences with the data service, the outcomes associated with NSDS outputs, and stakeholder attitudes toward the data service and related issues (e.g., data privacy);
- Emerging data sources, analytical methods, and computational techniques relevant to the NSDS; and
- How to design and launch education campaigns to address concerns surfaced via needs sensing and to highlight data service “wins.”

The NSDS should actively engage with the public to identify and answer questions of interest through, for example, search and discovery portals or a survey on evidence building.

## Function 3. Research and Development

The Committee agrees with the Evidence Commission on the importance of R&D for the NSDS and acknowledges that, unlike other data service functions, this role is not assigned to an existing agency within the federal statistical system.

### *Innovation Sandbox*

A key component of this R&D function is providing a neutral innovation sandbox for testing new and innovative technologies, software, and methods for risk-utility metrics, data sensitivity levels, access tiers, data protection protocols, and data analysis. For more information, see Recommendations Part 5. NSDS Technical Infrastructure and Tools—Innovation Sandbox.

### *Privacy-Preserving Technologies*

**Recommendation 3.12. The NSDS should promote the use of privacy-preserving technologies that support working with data in situ, coordinating with the research community to develop efficient, scalable tools for users from all levels of government (including through open competitions).**

The NSDS would be uniquely positioned to capitalize on the momentum around privacy-preserving technologies (PPTs) and to become a leader in PPT testing and adoption within the statistical space. As noted, federal agencies are leading the way in the development and deployment of PPTs—the role of the NSDS is to coordinate, support, and advance these efforts.

At a minimum, the NSDS should investigate ways to advance the PPTs explored by the multi-national group of statisticians and privacy experts within the [United Nations Global Working Group on Big Data](#): secure multiparty computation, fully homomorphic encryption, trusted execution environments, differential privacy, and zero knowledge proofs, most of which the Committee learned is already being explored by one or more statistical agencies. The NSDS could conduct this work internally and partner with outside agencies through its innovation sandbox. For more information, see Recommendations Part 5. NSDS Technical Infrastructure and Tools—Innovation Sandbox.

As part of this effort, the NSDS should help build capacity through training and technical support for users deploying these technologies. In addition, the NSDS should provide information to users and the public about evolving, state-of-the-art solutions for evidence building in a privacy-preserving, publicly auditable, and fair and equitable manner. For more information, see Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance above and Communication above.

The Committee is encouraged by ongoing PPT testing and implementation across the federal government, including at the Department of Education, NCHS, the Internal Revenue Service, and the National Institutes of Health (NIH); however, these efforts are not coordinated, making it difficult for agencies to share knowledge and leverage the work of others in this space.

The NSDS should find ways to build on, rather than duplicate, the fundamental research into PPTs that is already performed by other parts of the federal government, such as the statistical agencies, the National Science Foundation, NIH, and several projects sponsored by the Defense Advanced Research Projects Agency (e.g., [Brandeis, Data Protection in Virtual Environments, Programming Computation on Encrypted Data, SAFEWARE](#), and [Securing Information for Encrypted Verification and Evaluation](#)) and the Intelligence Advanced Research Projects Activity (e.g., [Automated Privacy Protection and Security and Privacy Assurance Research](#)). For more information, see the [ACDEB Project Inventory](#).

The NSDS should focus on demonstration projects that show what is possible, what fails, and why. The results of demonstration projects should be published, with a full accounting of labor and services, time and processor demands, and a description of where PPTs can promote or hinder the value that statistical services offer to society.

In addition, as a facilitator of the use of innovative methods, the NSDS may need to host the associated data until capability to support federated applications is broadly available. For more information, see Box 7. NSDS Data Hosting Services.

### **Box 7. NSDS Data Hosting Services**

As part of its coordinating and capacity-building functions, the Committee identified the potential need for the NSDS to host data. Data hosting is of particular importance when considering how the NSDS can provide services across the ecosystem, such as for state, territorial, local, and tribal governments that may otherwise lack the capacity and capability to host data. In its coordinating role, hosting data can facilitate efficiency and consistency for linked data sets and ensure quality and consistency of linkages that are resource-intensive and are expected to be needed again. The Committee recognizes cybersecurity risks associated with hosting data in a single repository and weighed such risks against the benefits of the services it can provide. To better ensure the reliability of linked data and facilitate verification and replication processes, it may be necessary for the NSDS to host linked data on occasion. The Committee discussed hosting data temporarily or only hosting the code used to construct linkages as possible alternatives.

For more information on PPTs and the related roles for the NSDS, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data Assets (Increase Adoption of Privacy-Preserving Technologies Within Government); Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance above; and Communication above.

## Function 4. Data Standardization

**Recommendation 3.13. The NSDS should coordinate with stakeholders to develop and promote standards for government data at all levels.**

As part of its coordination role, the NSDS should sponsor projects that encourage federal, state, and local partners to standardize metadata, file formats, and syntax. The results of these projects would increase the interoperability and usability of data for policy development and decrease cycle time for analysis. The NSDS should work with data providers, data users, and related communities of practice to refine these standards through iterative testing and piloting.

The Committee recognizes that data standards are an important component of improving data quality and promoting transparency and accountability. In its Year 1 report, the Committee acknowledged the importance of data standards, with nuances around data quality, common metadata, and data definitions. The value of data standards, consistency, and interoperability was echoed throughout the Committee’s fact-finding process and include the following examples:

- ***NCHS National Vital Statistics System (NVSS) Modernization.*** A major aspect of NVSS modernization is consistency across data sets, enabled by an Application Programming Interface (API) that supports the exchange of mortality data between NCHS and vital records offices. The API enhances data interoperability through timelier record level exchange and response, automation, and a more reliable and robust messaging platform. In addition, the API is built using the Fast Healthcare Interoperability Resources standard for the electronic exchange of health information that allows data to be reused by multiple parties for multiple purposes.
- ***The Midwest Collaborative (MWC).*** The MWC’s work to produce the Multi-State Postsecondary Dashboard has two primary goals: (1) combining administrative data from multiple states and (2) standardizing data models to reproduce measures across state lines. Thanks to the collaborative efforts of the states working together in the Applied Data Analytics training programs, the dashboard harnesses high-frequency and current data to provide intelligence for statewide and local stakeholders. The dashboard, enabled by access to a federally approved secure environment, has been adopted by several states and offers a scalable approach through which additional states could leverage such insights.
- ***Jobs and Employment Data Exchange (JEDx).*** [JEDx](#) is a data standards-based approach for how employers can produce enhanced and timelier data on both jobs and employment. The underlying hypothesis is that if agencies could align on a standard set of data to serve their purposes, then payroll processing companies could integrate those standards into payroll systems for both large and small businesses leading to more efficient reporting and higher quality data for analysis and decisionmaking.
- ***National COVID Cohort Collaborative.*** Recognizing that common data model mapping is critical, NCATS is currently working with different government agencies to create an accessible, dynamic repository of mappings covering data starting at its source in electronic health records through submission to the collaborative.

## Recommendations Part 4. NSDS Organizational Structure and Governance

In its Year 1 report, the Committee noted that it “seeks to advance the vision for how an NSDS fits alongside the federal evidence-building system and other key federal and non-federal actors to facilitate data access, enable data linkages, and develop privacy-enhancing techniques in support of increasing data availability for evidence building across the entire evidence-building ecosystem.”

In furtherance of that vision, the Committee outlined key attributes and functions to guide the establishment of the NSDS. In Year 2, the Committee continued to learn from ongoing efforts to implement NSDS functions under the Evidence Act and beyond—from previous experiences establishing national data services and large-scale research infrastructures, both domestically and internationally; from the expertise of its own members; and from outside experts who engaged with the Committee to explore how the NSDS could be designed.

Building on the evidence gathered over the past 2 years, the Committee proposes a governance structure that embodies the key attributes of the NSDS, as presented in the Year 1 report; advances the overarching purpose of the Committee to promote the use of data for evidence building; and effectively supports the data service’s core functions—coordination, communication, R&D, and data standardization.

The NSDS governance structure must be characterized by a clearly defined mission and clearly defined measures of success, which, as described by the [National Academy of Public Administration](#), are the first two principles for public organizations charged with improving policy outcomes. In addition, there must be meaningful mechanisms for diverse stakeholders to provide oversight and advice. Key considerations include the following:

- **Mission.** The NSDS has a complex mission to support the production of high-quality evidence while protecting privacy and confidentiality, consistent with the mission of a federal statistical agency. The NSDS must provide timely data access for technical and non-technical users, fulfilling the responsibilities described in CIPSEA 2018. These include advancing the twin goals of (1) increasing the value of data for evidence building through expanded access while (2) also ensuring the continued trust of data providers and the public that any access to data is designed to generate useful evidence, that privacy will be respected, and that confidentiality will be protected. To meet its mission, the NSDS must be established with stable workforces composed of highly trained technical talent who can objectively address long-term, complex problems.
- **Metrics.** In line with the primary responsibilities of statistical agencies codified in the Evidence Act and described in the Year 1 report, the NSDS must be fully transparent and accountable. This includes clearly communicating how and why it offers services. Likewise, NSDS performance must be measured through carefully selected metrics. If standards are not met, there must be clear mechanisms for addressing deficiencies and discontinuing or shifting services, if needed. The NSDS will have limited resources, and these metrics also provide one mechanism for prioritizing these resources.

- **Meaningful advice.** For the NSDS to be useful and used and to ensure that its functions supplement, not supplant, other evidence-building efforts and infrastructures, effective governance requires oversight and advice from diverse stakeholders. Key groups include government officials at the federal, state, territorial, local, and tribal levels; academic researchers; community organizations; and privacy, cybersecurity, legal, and ethics experts.

The organizational structure and governance for the NSDS must evolve over time and adapt to changes in technology, the legal environment, full implementation of CIPSEA 2018, public expectations, and lessons learned along the way. A phased implementation approach, informed by targeted pilot projects, is critical for standing up pieces of the structure and developing core functionality. This evolution must be predicated on a clearly defined theory of change. Box 8 outlines the major considerations for the NSDS Theory of Change. For more information, see the Supplemental Information—Other Models and Examples (Theory of Change) [posted](#) with this report.

## Box 8. NSDS Theory of Change

The NSDS theory of change can be broken into the five discrete categories shown in Figure A and described below.

**Figure A. The Theory of Change**



**Inputs.** Resources at the disposal of the project, including statistical and program staff, compute facilities, existing research, and budget.

**Activities.** Actions taken or work performed to convert inputs into outputs, specifically, the NSDS core functions—coordination, communication, research and development (R&D), and data standardization.

**Outputs.** The tangible goods and services that the project activities produce, including the following:

- Producing value for the American public while protecting privacy and confidentiality;
- Coordinating and supporting evidence-building efforts that cut across entities;
- Facilitating linkage of, secure access to, and analysis of nonpublic data;
- Providing capacity-building services for data users, data providers, and related communities of practice;
- Communicating the value and use of data for evidence building and how data are protected;
- Facilitating R&D and the adoption of practices and methods that enhance privacy and confidentiality and improve record linkage; and
- Fostering and promoting data standardization to enable more efficient and high-quality linkage, access, and analysis.

**Outcomes.** Results likely to be achieved once the beneficiary population uses the project outputs, including new evidence and products for decisionmaking, lower costs and higher quality evidence, greater transparency and accountability, communities of practice, and better collaboration across levels of government.

**Impacts.** The results achieved indicating whether project goals were met, including better decision-making; more timely, actionable, and policy-oriented research; policies that are more responsive to local conditions; and more effective local interventions.

In much of the discussion about evidence-based policymaking and the NSDS, the theory of change has been cut short, where the focus on inputs and activities stops with the endpoint of interest exclusively on outputs. The Committee encourages the OMB Director to work with agencies, the ICSP, the Evaluation Officer Council, the Chief Data Officer Council, and others to emphasize the value of short-term, intermediate, and long-term outcomes in the NSDS theory of change in relationship to the broader evidence ecosystem.

As part of its investigative process, the Committee explored a variety of organizational structures and governance models through the following mechanisms:

- ACDEB use cases on using administrative data to track project impact, education and workforce, health, labor market activity, and environmental quality and human health. For more information, see Appendix B. ACDEB Use Cases, Appendix E. ACDEB Subcommittee Guest Speakers, and the Supplemental Information [posted](#) with this report.
- Real-world emerging examples, including the America’s DataHub Consortium and the Standard Application Process (SAP). For more information, see Appendix D. ACDEB Virtual Site Visits and the Supplemental Information posted with this report.
- Reports authored by various stakeholders. For more information, see the appendix materials and the Supplemental Information posted with this report.
- An evaluation by the National Artificial Intelligence Research Resource Task Force on governance structures for large-scale research investments and operations. For more information, see the Supplemental Information—Other Models and Examples posted with this report.

This fact-finding process highlights that the federal government has many vehicles available for establishing organizational structures beyond standing up new agencies. These include entities that are owned and operated by the government, those that are owned by the government and operated by contractors, and federally funded research and development centers (FFRDCs), which are a special type of government-owned, contractor-operated structure. Vehicles for establishing organizational structures also include university or private-owned research centers. As noted in the Year 1 report, another possible vehicle is a public-private partnership.

In addition, multiple successful data infrastructure models exist across the private and public sectors. As the Committee learned from its health use case, the National Vital Statistics System modernization effort has benefited from a long-term and unique relationship with the [Alliance to Modernize Healthcare](#), an FFRDC. With its education and workforce use case, the Committee explored the Midwest Collaborative, which makes use of an administrative host and a data and training platform. There are also successful examples from the large-scale research environment. One size is not likely to fit all purposes.

In keeping with the Committee’s objective to provide advice that connects current activities and infrastructure to the target state for the NSDS, this section offers evidence, findings, and recommendations for establishing, exploring, and expanding the NSDS governance structure to realize the Committee’s vision for the data service and accelerate the promise of evidence-based policymaking across the data ecosystem.

## America’s DataHub Consortium as the Foundation for the NSDS

**Recommendation 4.1. NCSES, in coordination with OMB and the ICSP, should leverage Congressionally appropriated demonstration project funding to establish the America’s DataHub Consortium (ADC) as the pilot foundation for the NSDS. The ADC should sponsor pilots that demonstrate how data service structures and functions could grow, adapt, and evolve over time to realize the Committee’s vision for the NSDS.**

As noted, the organizational structure and governance for the NSDS must evolve over time, and a phased implementation approach is critical for standing up pieces of the structure. The Committee believes that the [ADC](#) offers a promising structure for that phased implementation approach. The ADC harnesses key characteristics of successful models and addresses the required organizational attributes and core functions for the NSDS. While the ADC in its current form does not offer the full set of structures, services, and functions envisioned by the Committee, it does provide a dynamic initial foundation for a fully realized NSDS.

Specifically, the ADC is an established entity operating under the auspices of the National Center for Science and Engineering Statistics (NCSES) with a vision that aligns closely with the aspirations for the NSDS. The ADC offers a flexible acquisition path that would allow the ecosystem to adapt, evolve, and ultimately offer the services and functions envisioned for the NSDS. The agility, speed, and potential scale of the ADC offer opportunities to experiment with mechanisms for delivering NSDS services in concert with other actors in the existing ecosystem. This could include submitting specific project or service requests for which ADC members could provide solutions and expanding services through contracted service providers or partners.

Furthermore, the ADC is uniquely positioned to take advantage of several opportunities in the evolving data ecosystem, including the following:

- The ADC sits under a CIPSEA-recognized statistical agency (that is, NCSES) with a broad mission to support evidence building.
- The [CHIPS and Science Act](#) authorizes funding for NCSES to “establish a demonstration project...to inform the full implementation of...a governmentwide data linkage and access infrastructure for statistical activities conducted for statistical purposes...”
- NCSES, under the guidance of OMB and the ICSP, is also overseeing the development of the SAP—the first data service function to be implemented, as envisioned by the Evidence Commission.
- The [President’s FY 2023 Budget](#) proposal includes funding for NCSES to “lead Government-wide development of evidence-building infrastructure, such as the Standard Application Process, America’s DataHub, and early work on a National Secure Data Service.”

Establishing the ADC as the pilot foundation for the NSDS would allow multiple initiatives, including the ADC, the SAP, and the NSDS, to develop in parallel and build upon one another in a way that supports a more seamless federal statistical system and, ultimately, harmonizes activities across the broader evidence-building ecosystem.

## NSDS as a Government-Owned and Contractor-Operated Organization

**Recommendation 4.2. The NSDS should be a legally recognized entity that is owned by the federal government and operated by a contractor.**

Under this approach, a dedicated government Program Management Office (PMO) staffed by permanent federal employees within NCSES should oversee NSDS operations with a dedicated contractor responsible for service delivery. The proposed NSDS organizational structure provides the flexibility to support the core NSDS functions, to innovate, and to recruit and pay high-quality staff, while also ensuring that the chosen contractor can be held accountable and replaced, if necessary. In addition, through contracted service providers and partnerships with government agencies at all levels, academia, and non-technical communities, the NSDS should introduce a multi-faceted approach to evidence building.

### *Diverse Stakeholder Oversight and Advice*

As noted above, to be successful, the NSDS should have meaningful oversight mechanisms for gathering advice from diverse stakeholders. Table 3. NSDS Oversight and Advisory Bodies outlines key considerations for identifying stakeholder groups and determining how they could provide timely oversight and advice.

**Table 3. NSDS Oversight and Advisory Bodies**

	Key Considerations
<b>People—the “who”</b>	Who are the key NSDS stakeholders to provide oversight and advice to the NSDS?
<b>Purpose—the “what”</b>	What is the purpose of the advice from each group?
<b>Placement—the “where”</b>	At what level of the organizational structure is the oversight and advice most effective?
<b>Power—the “how”</b>	From what legal authority do bodies derive their ability to provide oversight and advice to the NSDS?
<b>Phases—the “when”</b>	How to move from an initial structure to the target design?
<b>Priorities</b>	If advisory groups disagree, who decides the differences?

**People.** NSDS stakeholders fall into the following broad categories:

- **Policy setters.** Officials from government agencies at all levels must have meaningful input into the strategic direction of the NSDS to meet interdependent statutory responsibilities and to maintain control over data assets. Key stakeholders include the Chief Statistician of the United States; federal statistical agencies; other federal agencies, officials, and councils; and state and local governments.
- **User communities.** Stakeholders inside and outside the federal government must be able to guide NSDS operations to ensure that the NSDS is useful and used. Key groups include federal agencies, particularly lower capacity agencies and officials such as smaller statistical units, programmatic agencies, and evaluators; state and local governments; the research community; professional societies; and community organizations.
- **Technical advisors.** NSDS governance must include a mechanism for providing technical advice. Key stakeholders include privacy, confidentiality, cybersecurity, legal, and ethics experts from across the public and private sectors.

Together, these groups can provide comprehensive oversight of the NSDS and help ensure transparency and accountability. This is critical for the data service to meet its statutory requirements under the Evidence Act and to deliver coordination and capacity-building services to the broader evidence ecosystem.

**Purpose.** Stakeholders inside and outside the federal government should drive the strategic direction and key policy decisions of the NSDS. Similarly, for the NSDS to be a national resource, stakeholders must have direct input into NSDS operations and be able to provide timely, relevant technical advice. Different advisory bodies should serve different purposes—specifically, providing strategic direction, guiding operations, and giving technical advice—and should allow diverse communities to represent their unique perspectives and offer targeted recommendations.

**Placement.** The placement of the advisory bodies should directly relate to their purpose. Policy setters should provide strategic direction directly to NCSES through a Policy Steering Committee. User communities should guide operational decisions through a User Advisory Board. Technical experts should provide input through a Research and Technical Advisory Board. The User Advisory Board and the Research and Technical Advisory Board should advise the contractor responsible for the day-to-day NSDS operations—that is, the Operating Entity. In addition, there should be a Cybersecurity Advisor and a Privacy Advisor that report directly to NCSES.

**Power.** Non-federal representatives are legally restricted from voting on government issues. In addition, under applicable NCSES legal frameworks, stakeholders cannot advise contractors directly. That is, they cannot direct resource decisions (e.g., on staffing, spending, or scope of work). To harness the flexibilities of the ADC and allow the “right” stakeholders to provide the “right” advice to the “right” level of the organization would require a new legal mechanism.

**Phases.** Non-federal stakeholders could join the Policy Steering Committee as non-voting members. This would allow them to provide input into the strategic direction of the NSDS. In addition, the NSDS operating contract could be written with clauses that allow the Operating Entity to convene ad hoc working groups for targeted operational guidance and technical advice. The NCSES Program Management Office would direct the Operating Entity to implement these changes. Legal changes that would allow non-federal stakeholders to vote on policy issues or to advise contractors directly should be explored as part of the broader NSDS phased implementation approach. For more information, see Phased Implementation Approach below.

**Priorities.** Establishing separate advisory bodies could lead to divergent guidance and priorities. To ensure that these entities are coordinated, the Chief Statistician of the United States (or delegate) should serve as an ex-officio member of the User Advisory Board. In addition, in what the Committee envisions would be rare instances when advisory bodies disagree on a particular matter, the issue should be decided by the NCSES Director, representing the entity ultimately responsible for administering the funds to operate the NSDS.

### Phased Implementation Approach

To ensure the timely establishment of the NSDS and to start demonstrating value, the Committee recommends a phased approach for establishing the governance structure and standing up core functions. Table 4. NSDS Implementation Phases describes the primary phases—establish, explore, and expand.

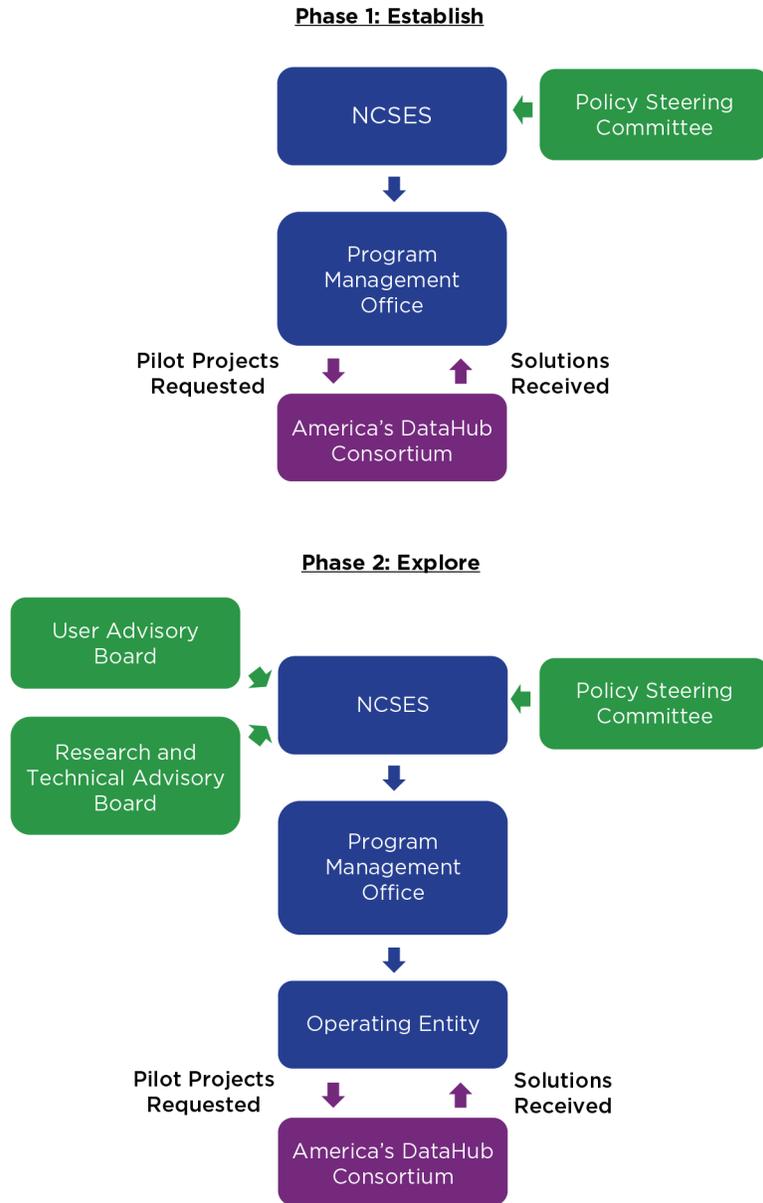
**Table 4. NSDS Implementation Phases**

Phases	Key Activities
Phase 1. Establish	<ul style="list-style-type: none"> <li>• With the ADC, NCSES should conduct pilots that test the main NSDS functions and that highlight the value associated with data linkage and data use.</li> <li>• NCSES should stand up the Policy Steering Committee, which should provide input into the selection and direction of the pilots.</li> <li>• NCSES should establish and staff the Program Management Office.</li> </ul>
Phase 2. Explore	<ul style="list-style-type: none"> <li>• NCSES should hire the Operating Entity and establish the User Advisory Board and the Research and Technical Advisory Board to guide operational and technical decisions. In this phase, the Cybersecurity Advisor and the Privacy Advisor could be members of the Research and Technical Advisory Board.</li> <li>• OMB, in coordination with NCSES and other stakeholders, should propose new legislation that would allow stakeholders to provide operational guidance and technical advice directly to the Operating Entity.</li> </ul>
Phase 3. Expand	<ul style="list-style-type: none"> <li>• With the appropriate legal mechanism in place, NCSES should shift the User Advisory Board and the Research and Technical Advisory Board to provide operational oversight and technical advice directly to the Operating Entity. The Cybersecurity Advisor and the Privacy Advisor should continue to provide guidance and oversight directly to NCSES.</li> <li>• The NSDS should transition to approving and operationalizing proposed projects, using service providers and partners, as appropriate.</li> </ul>

- ADC America’s DataHub Consortium
- NCSES National Center for Science and Engineering Statistics
- NSDS National Secure Data Service
- OMB Office of Management and Budget

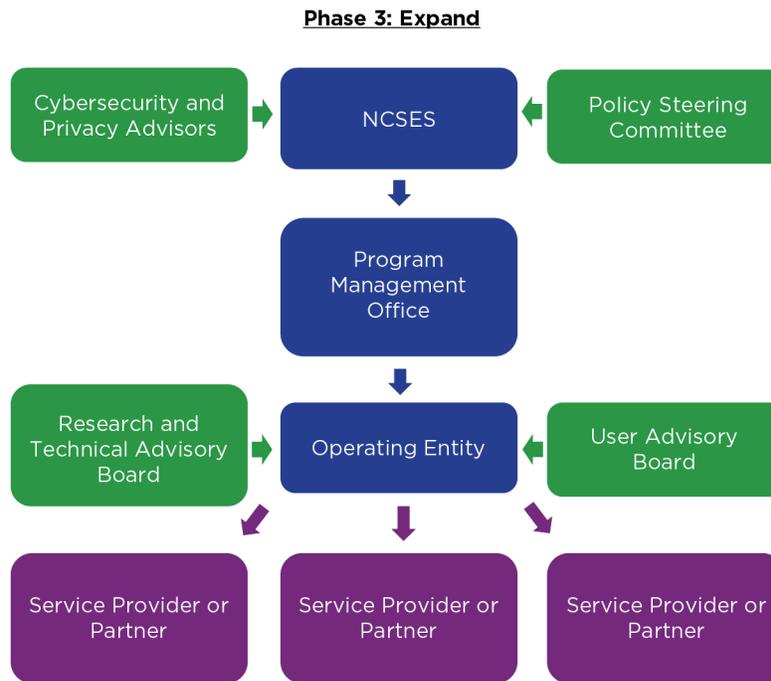
Figure 3 shows how the NSDS organizational structure could evolve.

**Figure 3. Phases of the NSDS Organizational Structure**



*Figure continues on next page*

**Figure 3. Phases of the NSDS Organizational Structure—*continued***



The Committee acknowledges that there may be other steppingstones to evolve the structure from the initial state to the target design. For example, NCSES may move from a project-based approach to more formal relationships with service providers and partners before shifting the User Advisory Board and the Research and Technical Advisory Board to the Operating Entity level—or vice versa. The path of evolution for the NSDS, including for the organizational structure and governance, depends on many factors, including sustainable funding, legislative changes, technological advances, buy-in and support from key stakeholders, incorporating lessons learned from the initial pilots, and other changes in the evolving evidence ecosystem.

Table 5 describes the components in the target NSDS organizational structure, including the composition, selection, legal and structural requirements, and roles and responsibilities. Note that the information presented in this table is illustrative, not comprehensive, and is meant to convey the strategic purpose and positioning of each component in the NSDS organizational structure, not necessarily to prescribe a specific implementation approach.

**Table 5. Components of the NSDS Target Organizational Structure**

Component	Composition and Selection	Legal and Structural Requirements	Roles and Responsibilities
<b>NCSES Program Management Office</b>	Includes the NCSES agency lead and supporting federal personnel, including staff with contracting expertise, as determined by the NCSES Director	Subject to the requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes and regulations	Oversees NSDS operations
<b>Operating Entity</b>	Staffed by experienced and qualified personnel, including chief security and privacy officers  Selected through a competitive process by the NCSES Director	Legally recognized entity with hardware, software, and an administrative infrastructure and capacity to support services  Subject to the requirements as established under contract with NCSES as well as requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes	Ensures effective delivery of NSDS services
<b>Service Providers and Partners</b>	Includes qualified organizations across the public and private sectors, such as statistical agencies, industry leaders, nonprofits, and academic institutions  May be selected by the NCSES Director or recommended by the appropriate advisory bodies, depending on the nature of the relationships and related requirements	Subject to the requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes and regulations as well as requirements established under contracts with NCSES, as appropriate  May require hardware, software, and technical infrastructure and capacity to provide select services	Provides select services, including targeted technical assistance on specific projects
<b>Cybersecurity Advisor</b>	Expert in cybersecurity threats and mitigation strategies	Subject to the requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes and regulations	Provides guidance on cybersecurity issues to other components of the NSDS organizational structure, as appropriate
<b>Privacy Advisor</b>	Expert in privacy and confidentiality	Subject to the requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes and regulations	Provides independent privacy and confidentiality oversight of NSDS activities with audit and investigative authority
<b>Policy Steering Committee</b>	Chaired by the Chief Statistician of the United States (or delegate)  Consists of federal Executive Branch stakeholders, such as the Chairs of the ICSP and other federal councils, including the Chief Data Officer Council, Evaluation Officer Council, Federal Privacy Council, and Chief Information Officer Council, as well as state and local government officials	Subject to the requirements of CIPSEA, the Privacy Act, the Paperwork Reduction Act, and all other applicable federal statutes and regulations	Provides strategic direction and advice on significant policy issues

Table continues

Component	Composition and Selection	Legal and Structural Requirements	Roles and Responsibilities
<p><b>User Advisory Board</b></p>	<p>Reflects the broader evidence-building ecosystem and includes members from federal agencies, state and local governments and their partner organizations, and the researcher and non-technical communities</p> <p>Members could be selected from a slate proposed by federal councils, professional societies, state and local agencies, and community organizations</p> <p>Chief Statistician of the United States (or delegate) serves as an ex-officio member</p>	<p>Subject to applicable legal requirements, including those established by new legislation</p>	<p>Provides operational guidance</p>
<p><b>Research and Technical Advisory Board</b></p>	<p>Includes privacy, security, legal, and ethics experts from across the public and private sectors</p> <p>Members could be selected from a slate proposed by federal councils, professional societies, state and local agencies, and community organizations</p>	<p>Subject to applicable legal requirements, including those established by new legislation</p>	<p>Provides technical advice</p>

CIPSEA Confidential Information Protection and Statistical Efficiency Act  
 ICSP Interagency Council on Statistical Policy  
 NCSES National Center for Science and Engineering Statistics  
 NSDS National Secure Data Service

## Holistic Governance Approach to Transparency and Accountability

**Recommendation 4.3. Through its governance and operations, the NSDS should model a holistic approach to transparency and accountability.**

Transparency and accountability are prominent themes that run through the Evidence Act. The NSDS must apply the CIPSEA framework, including meeting transparency requirements described in the forthcoming regulations on the responsibilities of statistical agencies and trust, expanding secure access to CIPSEA data assets, and the Standard Application Process (SAP). As part of its coordination and R&D functions, the NSDS should model best practices for implementing CIPSEA requirements and should catalyze innovation around these methods. For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance and Recommendations Part 3. NSDS Functions—Coordination (Advance Evidence Act Implementation for the Federal Government and Beyond).

Transparency and accountability are key attributes of the NSDS that must be built into business operations and flow through every component of the NSDS organizational structure and governance. Key elements of this holistic approach to transparency and accountability are described below.

**Key performance indicators (KPIs).** NSDS operational decisions must be driven by clearly articulated KPIs, including timelines and deliverables for the data service’s activities, outputs, outcomes, and impacts, consistent with the NSDS theory of change. These KPIs should include measures of data value, privacy loss, and jointly determined risk-utility implications of different access approaches, in line with requirements of the regulation on expanding secure access to CIPSEA data assets. For more information, see Box 8. NSDS Theory of Change and Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data Assets.

**Measure and report data value.** The production of value (or “utility”) is inherent to the core responsibilities of statistical agencies and, as such, is critical for the NSDS. There are several dimensions of value—broadly, adherence to democratic and equitable values and providing value to the public and, more specifically, value of the data assets, value of NSDS capabilities, and value of the data service itself. The NSDS should model an approach to measure and report on the value of each of these aspects, including the following actions:

- **Produce an NSDS data inventory with usage statistics.** The NSDS should develop and maintain a publicly available inventory of NSDS data assets in keeping with Evidence Act requirements for agency data inventories. While not a full measure of value, as a baseline, this inventory should include usage statistics. To support a more seamless experience for users, the NSDS data inventory should model the format and content, including detailed metadata, that could be used to harmonize other data inventories and catalogs.
- **Develop concrete measures of value.** The NSDS should develop and publish concrete measures of value, including exploring ways to measure the impact and the value of evidence for different stakeholders.

- **Implement the risk-utility framework.** The NSDS should explore different approaches, identify best practices, and model how to implement a risk-utility framework for federal agencies and the broader evidence ecosystem.
- **Report its work in a transparent manner.** The NSDS should publish an annual report that presents the results of these actions and should explore a dashboard to track metrics in closer to real time, building on the transparency metrics under development for the SAP.

**Measure and report privacy loss.** The NSDS should work with privacy experts in government, industry, and academia to develop measures of the privacy loss and harm associated with the applications of different privacy protections, particularly for vulnerable and marginalized populations. A report should be published annually.

**Jointly determine and report the risk-utility consequences of different access approaches.** The NSDS should jointly determine the value and privacy consequences of different access approaches, particularly for vulnerable and marginalized populations. A report should be produced annually.

**Independent external evaluation.** As an independent check on performance, the Operating Entity and the service providers and partners should be externally evaluated every 3 years by a professional evaluation organization. Key features of this evaluation include the following:

- The evaluation organization should be recommended by the NSDS oversight and advisory bodies and hired by NCSSES.
- The external evaluation should measure baseline KPIs, evaluate the results of the KPIs over time, assess the appropriateness of KPIs themselves, and compare the benefits of the data service to the costs.
- The evaluation results should be transmitted to the oversight and advisory bodies and then posted publicly.

For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data Assets, Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Standard Application Process, Recommendations Part 3. NSDS Functions—Coordination (Advance Evidence Act Implementation for the Federal Government and Beyond), and Recommendations Part 5. NSDS Technical Infrastructure and Tools—Accessibility (Harmonized Data Inventories and Catalogs).

## Near-Term Project Functions and Pilots

**Recommendation 4.4.** In its initial phases, NCSES, in coordination with the ICSP, should sponsor pilot projects that (1) explore the NSDS core functions—that is, coordination, communication, R&D, and data standardization, and (2) demonstrate the value of streamlining data sharing and coordination, specifically with projects that highlight cross-functional, cross-agency, and cross-governmental topics.

The Committee acknowledges the benefits of a phased implementation approach to establishing the organizational structure and governance for the NSDS. Furthermore, the Committee’s investigative process, especially the labor market activity use case, highlighted the value of pilots for building buy-in among key stakeholders, developing new products, establishing joint programs, and spurring best practices—all through a shared value proposition around better evidence for decisionmaking. For more information, see ACDEB’s labor market activity use case report in the Supplemental Information [posted](#) with this report.

Over the next year, a small number of pilots could be launched to test the robustness of the organizational approach and to develop initial KPIs. The initial framework should be used to identify the necessary inputs (funding levels), activities, outputs, outcomes, and impacts necessary to achieve the value propositions identified in the use cases and in the theory of change. The results of the pilots should be used to inform the final organizational structure of the NSDS. For more information, see Box 8. NSDS Theory of Change.

In addition, the chosen pilots should (1) advance efforts already underway, with an emphasis on topics related to the Committee’s use cases; (2) include federal agencies, states, and localities as well as private-sector researchers; and (3) clearly articulate the input and output privacy methods employed. Table 6 presents initial focus areas, grouped by NSDS core functions.

**Table 6. Near-Term Project Functions**

NSDS Function	Description
<b>Coordinate</b>	<p><b>Cross-functional, cross-agency, and cross-governmental linkage projects:</b> Support actors across the evidence ecosystem to link high-value data assets and leverage administrative data in new ways</p> <p><b>Technical assistance:</b> Stand up concierge services in line with a phased implementation approach</p>
<b>Communicate</b>	<p><b>Communications strategy:</b> Develop a comprehensive engagement and education plan, mapping out initial implementation steps</p>
<b>Research and development</b>	<p><b>Privacy-preserving technologies:</b> Advance research around privacy-preserving technologies and explore how techniques can be leveraged to expand “middle” access tiers</p>
<b>Data standardization</b>	<p><b>Data standards:</b> Develop and promote recommended standards for metadata, file format, and syntax for government data at all levels</p>

The next section presents additional details on three potential pilot projects.

## **Health Pilot**

### **Why?**

The COVID-19 pandemic has highlighted the need for more timely, accurate, and reliable health statistics at the federal and local levels. Decisionmakers are demanding better data to answer questions.

### **Desired Outcomes**

- What are current death counts, causes of death, and the level of “excess deaths”?
- How long is the average person expected to live, and how has this been impacted by the pandemic?
- How do death rates and causes of death vary by state? County? Race and Hispanic origin?
- What is the impact of COVID-19 infection during pregnancy on health services utilization and maternal and infant outcomes?

### **What would the pilot do?**

The pilot would explore the value of health data access, linkage, and analyses and demonstrate how the National Vital Statistics System could be enhanced to provide near-real-time information through better two-way communication and collaboration among federal and jurisdictional partners. Existing efforts by research teams across government, academia, and industry would be identified and assessed to determine how to scale their approaches. The pilot would also highlight the importance of standards, consistency, and interoperability across the data provided by health agencies at both the federal and state and local levels. Through expanded communities of practice, training, and resource sharing, the pilot would demonstrate the importance of robust governance structures to prioritize, execute, and monitor project goals and activities.

## ***Education and Workforce Pilot***

### **Why?**

Unprecedented changes in labor markets, as well as major federal infrastructure and technology investments, have led to fundamental changes in skill demands. Both sets of changes underscore the need to strengthen the connection between employment services, post-secondary programs, and workforce outcomes. Building these links, and understanding the impact of different measurement assumptions, will help individuals decide what education paths best meet their needs. The results will provide evidence to inform government decisionmaking at all levels about the highest return investments in skills that yield long-run economic security and mobility.

### **Desired Outcomes**

- Produce detailed local and national measures of the earnings and employment duration associated with post-secondary credentials;
- Provide information about the impact of training and credential programs provided in secondary and post-secondary settings;
- Provide information about the differences in earnings and jobs between those completing credentials and those leaving before completion;
- Strengthen the connection between employment services, post-secondary programs, and workforce outcomes;
- Characterize differences for different subgroups: race, ethnicity, sex, foreign-born, and first-generation status;
- Help individuals decide what education paths best meet their needs; and
- Encourage high-return investments in skills that yield long-run economic security and mobility.

### **What would the pilot do?**

The pilot would demonstrate the value of access to education and workforce data by state analysts and researchers. Analysts and researchers would be provided access to data to explore the impact of different linkage methodologies, privacy protection methodologies, and outcome measures on summary statistics for different subgroups. The use of training classes would be particularly encouraged so that federal and state agencies can test out different approaches, share knowledge, and build a knowledge base. The pilot would compare existing summary statistics with those produced using state and local data. The results would be shared through two-way communication and collaboration among federal and jurisdictional partners. The pilot would also highlight the importance of standards, consistency, and interoperability across the data provided by education and workforce agencies at both the federal and state and local levels. Federal and state agencies would collaborate to provide information about the value and applicability of different privacy protection methods for decisionmaking. Through expanded communities of practice, training, and resource sharing, the pilot would demonstrate the importance of robust governance structures to prioritize, execute, and monitor project goals and activities.

## ***Labor Market Activity Pilot***

### **Why?**

There is tremendous potential for administrative data on unemployment insurance benefits to enhance measures of labor market activity. The government response to the COVID-19 pandemic, and its effects on the economy and the employment situation for millions of workers, provides a recent example. New national and local measures could be developed to better describe the experiences of the unemployed, as well as the disparate impacts on different subgroups. These measures could include timely, actionable, and local information about unemployment duration, the number and frequency of employment and unemployment intervals, and the nature and quality of the reemployment experience. The results can support better evaluation and research leading to continuous improvement and program delivery. Local workforce boards could use the results to help inform resource allocation for unemployed workers.

### **Desired Outcomes**

- Help states better understand how different populations access unemployment benefits and identify potential barriers that those populations may encounter to receiving timely benefits;
- Improve the quality of data and fit for purpose;
- Identify opportunities to improve measurement, data collection, and data analysis;
- Improve program delivery by combining data across agency lines and creating new measures—on equity, on unemployment to reemployment, and on labor demand; and
- Improve data access, linking, and standards development.

### **What would the pilot do?**

The pilot would demonstrate the value of access to unemployment insurance compensation and wage record data by state analysts and researchers. Analysts and researchers would be provided access to confidential data to explore the impact of different linkage methodologies, privacy protection methodologies, and outcome measures on summary statistics for different subgroups. The use of training classes would be particularly encouraged so that federal and state agencies can test out different approaches, share knowledge, and build a knowledge base. The pilot would compare existing summary statistics with those produced using state and local data. The results would be shared through two-way communication and collaboration among federal and jurisdictional partners with workforce boards. The pilot would also highlight the importance of standards, consistency, and interoperability across the data provided by workforce agencies at both the federal and state and local levels. Federal and state agencies would collaborate with workforce boards to provide information about the value and applicability of different privacy protection methods for decisionmaking. Through expanded communities of practice, training, and resource sharing, the pilot would demonstrate the importance of robust governance structures to prioritize, execute, and monitor project goals and activities.

## Recommendations Part 5. NSDS Technical Infrastructure and Tools

The Committee has described ambitious and diverse activities for the NSDS that are designed to supplement, not replace, the activities of other actors in the ecosystem. To make the path to data for evidence building easier and more predictable—both for data access and privacy protection—the NSDS must provide technical infrastructure and tools to support its core functions, operating under the same CIPSEA framework as the federal statistical agencies. Technical infrastructure should address needs that include the following:

- Administrators need to manage an integrated inventory of all data available for use and linkage, as well as an auditable record of the provenance of which data sets were used for which evidence-building analyses.
- Administrators need systems that track the following: proposals, identity management, contracts with service providers, program management, agreements, policies, and information from advisory boards.
- Data contributors need tools to manage their inventories, requests, users, and outputs.
- Data contributors need assistance to create a policy governing the tradeoff between value and risk in order to control and potentially limit any disclosure or inference of their raw data. This policy might depend on the access tier of the user who wants to perform an analysis using their data.
- Data providers need the “plumbing” to acquire, handle, provision, compute on, and retain/destroy data, as well as a proposal management system that builds on the functionality of the Standard Application Process (SAP).
- Data managers and IT and security officers need auditors, penetration testers, and liaisons to obtain and maintain authority to operate.
- Data users need a point of entry; the ability to search and discover data leveraging the SAP; and methods and tools to review past projects, public information, and policies that pertain to their specific tier of access privileges.
- Data users need a proposal submission system; interfaces to access privacy protected data, which could include enclaves to access restricted data; web applications for creating and sharing computational tools such as Jupyter notebooks; and tools for implementing output privacy protections.
- Data users need to know the accuracy or confidence level of the results of their queries, validation servers to increase their own confidence in the results, and a standard way to reference data sets to encourage open and reproducible science.
- Data providers and users need web infrastructure to connect to an innovation sandbox and to engage in communities of practice and user groups with the aim of sharing knowledge that benefits the broader evidence ecosystem.

- Data providers and users need privacy-preserving technologies so that providers are comfortable approving data requests and users can gain data insights without access to the underlying data.

These needs are not unique to the NSDS, and they reinforce the importance of the forthcoming CIPSEA regulations and guidance and their related implementation. The CIPSEA regulations outline consistent processes, frameworks, and guidelines for meeting these needs. The remainder of this section presents recommendations for the technical infrastructure required to support NSDS operations and services. In many cases, these recommendations call out principles and best practices for the broader evidence-building ecosystem, especially the federal statistical system operating under CIPSEA.

## Accessibility

**Recommendation 5.1. The NSDS should provide technology so that users at any tier of access can safely and efficiently analyze data assets hosted by affiliated organizations, including federal, state, territorial, local, and tribal governments; nonprofits; and other organizations.**

As asserted in the Year 1 report and affirmed throughout this report, the NSDS should coordinate the use of existing infrastructure in the evidence ecosystem and provide capacity-building services that allow users to better access, link, and analyze data. Capacity building is key to ensuring equitable access to data by offering training, tools, and infrastructure to users who may not otherwise have access. To accomplish this, the NSDS must implement a robust technical infrastructure that meets Evidence Act requirements to expand secure access to CIPSEA data assets and lays groundwork for the future. The Five Safes framework provides an instructive model for designing such an infrastructure. For more information, see Box 3. The Five Safes Framework.

In addition, the Committee recognizes the importance of tools to make data discovery more accessible and to help users track applications through each phase of the approval process. As such, the Committee recommends the following:

**Recommendation 5.2. To support the discovery of data assets for evidence building, the NSDS should provide a technological process to support access to searchable and discoverable data, request data access, track the approval process, and document the outcomes of that process.**

The NSDS must follow the SAP, as mandated under CIPSEA 2018. In addition, that process should be integrated with the NSDS to meet the needs of data providers and users more seamlessly. This recommendation recognizes the need for technical infrastructure to support the proposed process and integration, including automated approaches to address user inquiries and publication of the knowledge base resulting from pilots and approved projects.

For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Standard Application Process and Harmonized Data Inventories and Catalogs below.

## *Harmonized Data Inventories and Catalogs*

**Recommendation 5.3. To support a seamless user experience, the NSDS should provide training and tools to harmonize the format and content across data inventories and catalogs and to ensure complete, consistent metadata are included in these inventories.**

To increase the discoverability of federal data assets, the Committee recommends that the NSDS should coordinate with the ICSP and the Chief Data Officer (CDO) Council to support a seamless experience for users accessing various federal data inventories and catalogs, including reducing burden for reporting agencies and minimizing redundancy. While each data inventory would be housed at the source agency, a consistent look and feel across the inventories with an efficient search capability would help users discover data sets more easily.

In addition, the NSDS should prioritize training and tools to help data providers develop and improve their inventories in line with the requirements of Titles II and III of the Evidence Act and work to date on the SAP data catalog. To support discovery, information about data sources should be:

- Curated to meet relevant criteria;
- Updated regularly;
- Made available to the public using methods that are easy to access and understand for both sophisticated and novice users;
- Feature usage statistics to track the uses of data assets and to serve as a possible baseline for measures of value. For more information, see Recommendations Part 4. NSDS Organizational Structure and Governance—Holistic Governance Approach to Transparency and Accountability; and
- Expanded, over time, to include information on any data asset used for an NSDS project, whether provided by the federal government, state and local governments, or the private sector.

The Committee emphasizes the importance of data discovery in the evidence-building process and recognizes that transparent, clearly documented, publicly available agency data inventories and integrated data catalogs are crucial for supporting data discovery. In the Year 1 report, the Technical Infrastructure subcommittee asserted that there is a role for the NSDS to play in this space by recommending that “the NSDS data concierge service should coordinate with the Federal CDO Council and statistical agencies to advance efforts to index data inventories and metadata repositories.”

Within the federal government, each agency is required to have a data inventory, and there are several data catalogs to inform users of possible data sources. Descriptions of two primary examples of such catalogs—data.gov and the SAP data catalog—are included below.

- **Data.gov.** Created in 2009, [data.gov](https://data.gov) is the first national data catalog. Under a 2013 OMB policy, agencies are required to contribute entries from their comprehensive data inventories to this catalog. Title II of the Evidence Act, or the OPEN Government Data Act, requires each agency to develop and maintain a comprehensive data inventory and accompanying metadata for each data asset. These data assets and metadata are to be made available through the federal data catalog. Data.gov, maintained by the General Services Administration, currently serves that function.

- **SAP data catalog.** The ICSP is building the SAP data catalog to help users explore what restricted data they might apply to use across the holdings of all participating agencies. In addition to this more tailored scope, this catalog provides greater “discovery” metadata than is available via data.gov. The catalog links to the websites of agencies that provide access to data assets and additional metadata rather than replicating the information in two places.

While these catalogs meet basic legal requirements and provide building blocks for future expansions and integration, there is much room for improvement, coordination, and harmonization. For instance, as noted in the Year 1 report and reinforced by Committee discussions during Year 2, data.gov needs curation and quality control to increase the value of the site for users. Likewise, the SAP catalog could be expanded to include helpful information, such as indicating which files are linkable.

## *Searchable Project Inventory*

### **Recommendation 5.4. The NSDS should collect and house a searchable inventory of projects that highlights what data sets are being used for what purposes.**

The Committee agrees with the Evidence Commission recommendation 4-3: “To ensure exemplary transparency and accountability for the Federal government’s use of data for evidence building, the NSDS should maintain a searchable inventory of approved projects using confidential data and undergo regular auditing of compliance with rules governing privacy, confidentiality, and access” (that is, a “transparency portal”). Congress enacted this recommendation as part of establishing the SAP.

In addition, a major barrier to better evidence building and broader collaboration is that organizations and researchers may simply not know about the related work of other groups. There have, however, been recent strides in leveraging advanced technologies to gather and compile this information. For example, as described in the Year 1 report, the Coleridge Initiative’s [Show US the Data](#) competition explored how data science techniques, like machine learning and natural language processing, could be utilized to discover how public data are used and by whom. Based on the winning models, the Coleridge Initiative began developing several proofs of concept, including data usage scorecards and automated data inventories. This work has transitioned to New York University. In addition, the Census Bureau regularly releases an inventory of projects that use one or more Census Bureau data assets through the Federal Statistical Research Data Center network. This list is indexed by data set, investigator, research location, and date and will be integrated with publications resulting from the projects.

These efforts to understand the “rich context” of research can be used to answer questions like: How many publications use the data set? How is that use changing over time? How does that compare to other data sets produced by the agency? What topics are the data set being used to study? What are the publications associated with each topic? Who are the authors using the data set? What are the journals publishing work using the data set? Such approaches could be adopted for a broader audience, supporting knowledge transmission among state and local evidence builders in addition to federal and academic researchers.

Another challenge is that researchers may not use standard formats to cite their work. The use of tags, such as Digital Object Identifiers (DOIs), could make it easier to gather information on which data sets are used for which types of projects and could feed machine learning models to automate the compilation of such information.

## Data Integrity

**Recommendation 5.5. The NSDS should provide tools and support to users in conducting secure, accurate, and scalable analyses. In addition, the NSDS should facilitate the development and deployment of data access protocols that offer alternatives to the standard direct data access models.**

As part of implementing the CIPSEA requirements to expand secure access to federal government data assets, the NSDS should support the introduction of new tiers of access and efforts that make existing tiers more efficient. Whether users seek summary tabulations, dashboards, regression analyses, or machine learning models, the NSDS could offer services to support these activities. These may include training or connections to peers addressing similar issues or the use of technical infrastructure. For example, the NSDS could support agencies exploring synthetic data with verification or validation servers to ensure the integrity of results. The NSDS could also introduce data providers to artificial intelligence approaches that can automate some data cleaning steps, allowing them to move away from pre-packaged, rule-based processes and make more data available for evidence building.

Agencies where data handling is more mature, such as the 13 primary statistical agencies and units for which statistical activities are a major part of their missions, are in a better position to adopt the new tiered access framework as required under CIPSEA 2018. In contrast, agencies that only focus on data processing for administrative tasks and that may not currently support researchers and evaluators will face demands for metadata, data structures, and efficient provisioning that they have not faced before. The NSDS, in its coordination and capacity-building role, should help meet these needs.

For more information, see Recommendations Part 3. NSDS Functions—Coordination (Support a High-Quality User Experience, Including Providing Technical Assistance) and Recommendations Part 3. NSDS Functions—Research and Development.

## Disclosure Limitation

### **Recommendation 5.6. The NSDS should facilitate the development and application of statistical disclosure limitation methods.**

The Committee recognizes that the NSDS should help advance disclosure limitation techniques, including investing in more open-source tools and training, encouraging more researchers to contribute to this work, and aiding users with varying levels of technical expertise. This activity should be supported by infrastructure to host tools and training included in the Federal Committee on Statistical Methodology’s (FCSM) [Data Protection Toolkit](#) and should feature an active research program to develop new approaches and training.

To ensure that methods applied are developed using realistic risk models, the NSDS should provide users with infrastructure and support for conducting comprehensive risk assessments as required by CIPSEA 2018. The NSDS should facilitate the completion of privacy risk assessments for its users, including implementing standards to conduct such assessments outlined by the related CIPSEA regulation, providing technical assistance to identify appropriate tools and mitigation strategies, and maintaining information on privacy risk assessments from past projects that can be integrated into the federal data catalog.

For more information on disclosure limitation and the role of the NSDS, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance, Recommendations Part 3. NSDS Functions—Coordination (Advance Evidence Act Implementation for the Federal Government and Beyond), and Recommendations Part 3. NSDS Functions—Coordination (Support a High-Quality User Experience, Including Providing Technical Assistance).

## Auditability

**Recommendation 5.7. To ensure auditability and transparency, the NSDS should provide tools to track the provenance of all data sets and all evidence-building analyses performed over those data sets.**

This should include tools for assessing and reporting on the quality of data assets, supporting data validation, and facilitating reproducibility.

**Data quality.** The Committee recognizes the importance of data quality and the challenges of ensuring fitness for use, especially with administrative data sets that have not been traditionally used for evidence building. As stated by the Technical Infrastructure subcommittee in the Year 1 report: “NSDS must support fitness for use assessments and communications. To accomplish this, there must be an investment in the NSDS infrastructure that allows for both assessing and publicly communicating the quality of data prior to and after evidence building. In terms of data capacity and readiness, huge variation exists across agencies at the federal, state, and local levels. The NSDS must meet data providers where they are, providing a secure environment and tools to assess and monitor data quality and to report out on data fitness for use and planned uses.” The FCSM’s [“A Framework for Data Quality”](#) and subsequent case studies provide a useful foundation for helping agencies and their data users assess data quality.

**Data validation and reproducibility.** There are several ways that the NSDS could support data validation and facilitate reproducibility, including the following:

- Data validation, especially for those using synthetic data, is an important service. For example, for linked data sets available on validation servers coordinated by the NSDS, it may make sense for the data to be hosted by the NSDS for efficiency and to ensure consistency in security.
- For restricted-access data sets, the NSDS could provide as a service the ability to attest to the origin of data sources and the analytical techniques applied to them to increase confidence in scientific findings that cannot be easily replicated.
- The NSDS, in coordination with the ICSP and the CDO Council, could develop a standard way to reference data sets that encourages and facilitates reproducible science. Such a standard could be adopted for agency-level data inventories and the federal data catalog.

OMB should support such efforts more broadly by issuing specific guidance, building on the [recent National Academies of Science, Engineering, and Medicine study](#) that examines ways to enhance transparency and reproducibility of agencies’ statistics and to improve statistical program workflow processes.

## Data Linkage

**Recommendation 5.8. The NSDS should provide tools and support to users in conducting scalable, privacy-preserving record linkages, facilitating data preparation and review of matching metrics. As part of its data concierge services, the NSDS should coordinate with federal, state, territorial, local, and tribal government officials seeking linkage services.**

When users want to learn about common patterns or features between two or more data sets, and an existing statistical agency or partner is not already prepared to meet that need, the NSDS could assist with data linkages in support of distinct authorized projects.

Whenever possible, the NSDS should use privacy-preserving record linkage (PPRL) technology, which enables finding and analyzing links between data without the need for data sharing (either with each other or with the NSDS). The Committee hosted several external experts who spoke about PPRL, including Actuate Innovation, Census xD, Datavant, the National COVID Cohort Collaborative, Opportunity Insights, and the Urban Institute. The work of these experts shows that PPRL technology can be used in a variety of evidence-gathering applications. For more information, see the appendices and the Supplemental Information [posted](#) with this report.

For record linkages that are beyond the abilities of current PPRL technology, the NSDS may be able to connect some users to existing linkage programs at statistical agencies (e.g., the Census Bureau and the National Center for Health Statistics). For other users, the NSDS could act as a facilitator for temporarily storing some features of the data sets. This work should be “as transient as possible,” so the NSDS could assist data owners by connecting them to a secure place for joining data and running analyses. In addition, to support linkages across data providers who may lack capacity and capabilities (for example, state and local governments), the NSDS could host input data to ensure quality and consistency in linkages that are resource-intensive and are expected to be needed again. Any data stored by the NSDS should follow best practices for data protection at rest and in transit and should promptly abide by any data deletion request by the data owner. For more information on the role the NSDS could play in hosting data, see Box 7. NSDS Data Hosting Services.

The Committee explored several federally sponsored and approved secure environments, including the FSRDC network and the Federal Risk and Authorization Management Program. In addition, some states have internal service providers, located in government and in public universities, who directly access state administrative data systems to extract data for linkages and analyses on their secure data systems. Other states or localities may choose to use the NSDS linkage services more intensively.

For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data (Tiered Access in Practice), Recommendations Part 3. NSDS Functions—Coordination (Support a High-Quality User Experience, Including Providing Technical Assistance), and Recommendations Part 3. NSDS Functions—Research and Development.

## Innovation Sandbox

**Recommendation 5.9. The NSDS should feature a sandbox for testing new and innovative technologies and software for multiple data access tiers, data protection protocols, and data analysis.**

The NSDS should create secure spaces where new products and approaches could be developed and tested. The NSDS should be a neutral ground between agencies where secure testing of the quality of new data linkages (particularly the impact on population subgroups), privacy-preserving technologies, training, and product development can occur. During its investigative process, the Committee discussed how data providers and users throughout the public and private sectors could benefit from such a capability. The NSDS should create a sandbox for use by government agencies, corporations, academic researchers, and nonprofit think tanks and foundations. The sandbox should support the NSDS R&D function and efforts under the proposed federal statistical system learning agenda. For more information, see Recommendations Part 2. Other Evidence Act Items—Federal Statistical System Learning Agenda and Recommendations Part 3. NSDS Functions—Research and Development.

The sandbox would allow the broader evidence-building community to understand specific challenges and to determine whether technologies meet the users' needs. For example, the Statistics of Income (SOI) Division at the Internal Revenue Service has aggregated data for its website, but these data are not yet included in the SOI Application Programming Interface. The sandbox could serve as a test bed to compare the risk and utility tradeoffs of current statistical disclosure limitation techniques with other approaches, especially those involving open differential privacy approaches and model-driven methods to set and monitor privacy budgets. Moreover, the sandbox could sponsor competitions, such as re-identification challenges, to stress-test any proposed synthetic data generators, privacy-aware data synthesizers, anonymizers, diagnostics, and more.

The goal of this sandbox is not to lock into a current solution, or on the other extreme, to wait passively for the future to happen. Instead, the sandbox should be a laboratory for: (1) testing and improving new technologies and how to incorporate them, (2) developing new products, and (3) providing a pathway for technology transition of mature products into the NSDS. Box 9. Technologies Ripe for Innovation Within the Sandbox provides examples of how the sandbox could be used.

## Box 9. Technologies Ripe for Innovation Within the Sandbox

Promising new technologies, such as synthetic data, validation servers, and secure multiparty computation (SMC), could allow significant expansion in research use of administrative data while providing more robust privacy guarantees than traditional disclosure control methods. However, these technologies are still in relatively early stages of development.

Regarding output privacy and disclosure limitation, one issue is measuring the privacy risk associated with *any* data release. Differential privacy (DP) is one approach; it provides a strong level of output privacy in that it protects against both known and unknown threats to the data. But the DP benchmark has only been applied to a limited class of data releases, and, moreover, it is a topic of active debate to determine which applications require its strong privacy guarantees versus where more targeted protections might be possible. Developing more flexible relaxations to DP that are consistent with real-world threats should be a research priority.

Developing high-quality synthetic data that are consistent with a formal privacy standard and helping data stewards manage the tradeoff between privacy and utility would also help both data providers and users. Similarly, validation servers designed to handle a wide range of statistical analyses that are consistent with a formal privacy standard could significantly increase the research utility of sensitive data, but there are many unanswered questions about how to implement those servers and how to measure and manage a privacy budget—a measure of cumulative privacy loss across many queries. This has been done for particular statistical models in the context of DP but not more practical variants.

Regarding input privacy and protection of sensitive source data, it has already been discussed above that linking data sets is enormously valuable for research. For example, drawing program information from one data set and demographic information from another could allow researchers to study the disparate effects of policy interventions by race and ethnicity, or education level. SMC allows researchers to conduct certain kinds of statistical analysis using data from multiple data sets without ever formally merging the data or even having the data sets leave the control of their original owners. This technology has been a topic of research for 40 years, with substantial funding for fundamental research by the National Science Foundation, the Defense Advanced Research Projects Agency, the Intelligence Advanced Research Projects Activity, the National Institutes of Health, and other basic science funding agencies. The NSDS innovation sandbox would provide the SMC community with a better understanding of the specific needs of statistical agencies, so they can customize and optimize SMC tools to provide the most value to NSDS stakeholders.

Finally, these technologies are new and unfamiliar to both data owners and users. Providing resources to help data stewards apply the new technologies and help users to understand the technologies and, in some cases, new ways of conducting research would help ensure that the new technologies are accepted and used appropriately.

## Recommendations Part 6. Resources and Funding

As the Committee discussed resources and funding, ideas coalesced around two interrelated sets of resource needs. Those include the following:

- Resources to advance evidence-based decisionmaking throughout the ecosystem, even without the introduction of the NSDS; and
- Resources for the NSDS to meet its mission as a specific intermediary, supporting data providers and users.

The Committee asserts that, for the data ecosystem to flourish, there must be adequate resources available to both (1) data and data service providers across all levels of government and (2) public and private-sector data users, using mechanisms that are transparent, equitable, and sustainable. If any one segment of the data ecosystem is not “fed” or “nurtured” then, ultimately, all parts of the ecosystem will be impacted. As a coordinator of services across the broad evidence-building ecosystem, the NSDS, as envisioned by the Committee, is strategically positioned to use its own resources to help meet the needs of both data service providers and data users in collaboration with existing actors in the ecosystem.

The recommendations in this section outline an ambitious, far-reaching plan for resourcing the evidence-building ecosystem. Many of these recommendations call on OMB to take action related to the FY 2024 Budget. The Committee strongly encourages the OMB Director to implement these recommendations in the given timeframes. The Committee acknowledges that the size and impact of these recommendations may require coordination and planning that could extend action into the FY 2025 Budget cycle.

### Resources to Enhance the Existing Evidence-Building Ecosystem

In keeping with the Year 1 report, the Committee concludes that “The NSDS will not succeed without proportionate investment in data producers at all levels, including federal, state, and local partners.” Governments across the evidence ecosystem also need resources to support them as *users* of evidence for decisionmaking. Considerations and recommendations around resources for federal agencies and for state, territorial, local, and tribal governments as data providers and users, even without the existence of the NSDS, are described below.

#### *Federal Agencies’ Implementation of the Evidence Act*

In the Year 1 report, the Legislation and Regulations focus area noted that “for the most part, [federal] agencies have not received supplemental appropriations or necessary funding flexibilities to support implementation of the Evidence Act and Federal Data Strategy” and recommended that, “in addition to any direct appropriations, the OMB Director should propose legislative flexibilities for facilitating funding set-asides for data infrastructure and analysis activities, recognizing these activities are core functions of government.”

During its second year, the Committee continued to explore the needs of both statistical and nonstatistical agencies in implementing the provisions of the Evidence Act, pulling in the expertise of its own members. The Committee’s discussion of these needs falls into two general categories—support for Evidence Act officials and resources for federal statistical agencies to advance evidence-building activities across the data ecosystem.

First, while the Evidence Act established new roles (i.e., Chief Data Officers (CDOs), Evaluation Officers (EOs), and Statistical Officials (SOs)), responsibilities, and frameworks, the Committee recognizes that federal agencies could benefit from having additional staff and resources to support these officials and to advance evidence-building activities, like evaluation. As is clear from federal budget proposals and third-party surveys of the [Chief Data Officer](#) and [Evaluation Officer communities](#), there is a great deal of heterogeneity in the staffing and resourcing of these functions today across the federal government. Yet, there is clearly a great need for additional resources and capacity to achieve the expectations that were outlined in the Evidence Act when enacted in 2019.

Likewise, federal statistical agencies need more personnel to review applications for data access more quickly and to provide enhanced technical assistance to users of their public and confidential data assets. For example, even a sophisticated evaluator may need guidance and assistance from statistical agency experts on how best to leverage government data.

In addition, federal agencies need additional resources to help cover the costs of developing documentation, data cleansing, and formatting data for analysis that support the targeted projects and activities discussed in this report. The Committee recognizes that these activities are time consuming and costly parts of the process of creating evidence for decisionmaking. Many of the outside experts, including those from the National Institutes of Health Library Data Services, Federal Statistical Research Data Centers, the National Center for Advancing Translational Sciences, and the Inter-university Consortium for Political and Social Research, echoed this need in the context of both technical infrastructure and technical assistance. Agencies could be funded for the cost and expertise needed for this process, targeting data sets for which there is high demand and a large value-add. Incentivizing the agency that is responsible for the data to prepare them for use could spare time and expense for users who will ultimately leverage the data. For more information on usage statistics and value metrics to help agencies make determinations of which data assets to target, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data Assets (Risk and Utility Framework).

The federal government cannot fully realize the goals of the Evidence Act without additional resources for the ecosystem, specifically to support CDOs, EOs, SOs, and the federal statistical system.

## Near-Term Resource Strategy for Evidence Act Implementation

**Recommendation 6.1. The OMB Director and Agency Heads, in consultation with the designated Chief Data Officer, Evaluation Officer, and Statistical Official at each agency, shall allocate funds from appropriations to adequately resource and support evidence-building activities for FY 2023. In addition, the OMB Director should prioritize direct appropriations and funding flexibilities as part of the FY 2024 Budget formulation process and encourage all Agency Heads to prioritize Evidence Act implementation activities going forward.**

OMB should be directing agencies to allocate resources to implement the Evidence Act and related administration priorities, like those described in the President's [Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking](#) to Agency Heads (January 27, 2021). Because every agency has different needs, it is important for the officials responsible for this implementation to be involved in resource allocation decisions and the budget formulation process.

In the Year 1 report, the Committee acknowledged the need for the OMB Director to prioritize resources for the Evidence Act in the President's FY 2023 Budget request to Congress. While OMB went to great lengths to recognize the President's and the Administration's commitment to evidence-based policymaking in the Budget documentation, major investments for the capacity of government and specifically for CDOs, EOs, SOs, and the federal statistical system were not adequately prioritized in that budget request. The Committee is disappointed that OMB did not provide more direct appropriations, funding flexibilities, or set-aside authorities to support Evidence Act implementation in FY 2023. This is an area that OMB should remedy in the years ahead by taking a more systematic approach that ties resource requirements to other standards and expectations of the Evidence Act.

## Systematic Resource Planning for Evidence Act Implementation

**Recommendation 6.2. OMB, in coordination with the ICSP, the CDO Council, the Evaluation Officer Council, and other relevant federal councils, should develop a systematic approach to funding Evidence Act implementation.**

While the Committee’s charge is intended to focus in part on the eventual creation of the NSDS, it is primarily tasked with focusing on aspects of OMB’s implementation of the Evidence Act and related activities. As emphasized throughout this report, the Committee believes that the Evidence Act holds much promise for reshaping decisionmaking across the whole of government, touching every program in every agency. This requires the federal government to take a different approach to resources for evidence building with a focus on planning and allocation, not just appropriations. Agencies must plan resources carefully for data collection, data management, and data governance, including how they will securely expand access to data assets. Adequate consideration of resourcing and how to promote transparent processes that hold officials responsible for budgeting for those activities accountable is both justifiable and necessary.

For evidence-building activities to be a priority and for the vision of the Evidence Act to be realized across government, including supporting evidence building as a shared function of existing resources, government agencies should increase transparency about their resource plans, needs, and spending (or lack of) for building and using evidence.

A hallmark of the Evidence Act is that evidence building requires systematic planning, and the law itself has several touchpoints for expanding this vision to support an approach that better surfaces and funds agency needs. The Committee discussed several options for (1) formalizing the planning process for evidence-building resources, (2) increasing transparency in the budgeting process, and (3) identifying the resource gap to support Evidence Act implementation. Highlights of these discussions are described below.

**Strategic planning process.** The strategic planning process may provide an opportunity for agencies to be more explicit about the resources needed to implement Evidence Act requirements that would help raise the visibility of funding issues. By design, this process is more transparent than other processes—like the budget formulation process—and requires quarterly performance reviews. Agencies already produce learning agendas and evaluation plans, so it could be a natural extension of the process to discuss the budget resources for CDOs, EOs, and SOs to support these plans. Strategic plans, however, tend to be presented at a relatively high level that does not include resource requirements. In addition, while final plans and learning agendas are public, there is not much visibility into the deliberative process to develop these documents.

**Capacity assessments.** Under the Evidence Act, each agency’s strategic plan is required to include “an assessment of the coverage, quality, methods, effectiveness, and independence of the statistics, evaluation, research, and analysis efforts of the agency.” These capacity assessments could be used as a vehicle for enhanced resource planning; however, for this to be successful, agencies may need additional guidance around how to produce these assessments, including standard assumptions and criteria.

**Agency budget requests.** Traditionally, budget requests submitted by agencies are withheld by OMB as deliberative and not deemed public record under the Freedom of Information Act. OMB required agencies to submit separate information in the deliberative budget process for data and evidence priorities in consideration of the FY 2024 Budget that will be published in spring 2023. The Committee discussed the possibility of OMB creating a separate class of budget submission for data and evidence-related activities that could be published publicly, similar to the information OMB already requests from agencies for its own deliberations. This would allow OMB to make that information available to the public and to demonstrate its prioritization and commitment to funding these items. Certain states take a different approach to transparency in the budget process. For example, in Missouri's state budget process, the budget tables provided to the general assembly highlight differences between departmental budget requests submitted to the governor and those submitted by the governor to the general assembly.

While there is a great need for transparency around evidence-building needs and related funding requests, opening agency and departmental budget formulation processes for public review has several downsides. First, the current process allows administrations to speak with a unified voice and prioritize initiatives across departments. The process is already time-consuming, and this would likely insert another level of pre-decisional meetings. Typically, CDOs, EOs, and SOs provide some input into budget requests. An unintended consequence of releasing pre-decisional information could be that agency heads would limit the flow of ideas during the process. Finally, even if agencies provide explicit information on their budget proposals related to evidence-building activities, this does not mean that those initiatives will be funded, so another round of assessment would be required once Congress passes the ultimate budget.

**Inclusion in the budget process.** The Committee explored [OMB Circular No. A-11](#) as a helpful model for strengthening the role of officials tasked with implementing the Evidence Act (like CDOs, EOs, and privacy officials) in the budget process. Specifically, when an agency wants to procure information technology that would process personally identifiable information, the agency privacy official must review the proposed acquisition and determine costs from a privacy perspective. Similarly, Chief Information Officers have authority to approve IT requests with the long-term vision of building a coherent IT infrastructure.

A clearly defined and well-documented approach that includes officials responsible for evidence-building activities in the budget process could improve long-term planning and resource allocations. Hearing from more stakeholders in the budget process would be more inclusive; however, the Committee also discussed challenges and potential roadblocks with this approach. First, it could add another layer to an already bureaucratic budgeting process. In addition, because the organizational positions, roles, and influence of CDOs and EOs vary greatly across agencies, the impact of their involvement in the budget process would also vary.

**External evaluation.** Agencies' evidence-building needs and the related funding levels could be assessed periodically by a third-party objective reviewer to ensure alignment and highlight gaps. Such an assessment could include minimum viable resource levels for required activities for CDOs and EOs, for example, based on criteria that reflect OMB guidance.

The success of Evidence Act implementation comes down to changing the mindset around these activities by demonstrating the value of evidence-based policymaking. Once agencies see their own value proposition, they will be considerably more likely to prioritize these activities in their budget processes and plan for them accordingly. Systematic, transparent resource planning provides one mechanism for helping agencies discover this value. This approach goes hand in hand with broader coordination efforts around Evidence Act implementation.

## Designated Funding Stream for Evidence-Building Activities

### **Recommendation 6.3. The OMB Director should request funding and authority in the FY 2024 Budget for a new interagency budget account with transfer authority to support Evidence Act implementation.**

Increasingly, agencies have shared or cross-agency evidence and data needs or priorities. The Administration recently demonstrated this with the publication of a cross-government learning agenda under the President’s Management Agenda. Funding these types of initiatives can be challenging with mixed incentives and funding restrictions, subject to limitations under individual agency appropriations. A shared interagency fund with resources to support data management, collection, and analytical priorities can help mitigate such limitations while providing supplemental resources for top priorities. Specifically, allocating a new interagency budget account to OMB would also align the resources to emerging cross-agency priorities jointly identified through the interagency learning agenda process, for example.

The Committee discussed options for a designated funding stream, including the following:

- ***Technology Modernization Fund (TMF)***. The TMF, established by the [Modernizing Government Technology Act of 2017](#), provides funding for agency IT modernization initiatives. The fund has received an annual appropriation of \$175 million, with an additional \$1 billion as part of the 2021 American Rescue Plan. To access these funds, agencies must send proposals to a Technology Modernization Board and, once projects are approved, must provide detailed project spending plans highlighting key milestones. There could be a similarly designed fund established to support Evidence Act implementation.
- ***Evidence Incentive Fund***. The Data Foundation presented its vision for a shared fund for Evidence Act implementation. The proposed Evidence Incentive Fund “would serve as a funding vehicle that incentivizes activities to implement the Evidence Act, in combination with appropriate transparency and engagement on key provisions that support accountability and oversight of the initiatives.” Such a fund was initially recommended by the Evidence Commission and envisioned to function like a working capital fund with authorized funds for agencies creating learning agendas. The fund would be built by “rolling over some of an agency’s unobligated balances at the end of a fiscal year, then making the resources available without a time limitation for use.”

In practice, elements of both models are likely relevant for the interagency fund to become operational. The TMF uses an application approach. The Committee recognizes that the capacity gaps are currently vast across government; thus, the fund could receive applications for filling gaps and select the most critical needs to fill in the current year. Similarly, the Evidence Incentive Fund leverages authority to capture unused resources that would otherwise be unavailable to agencies and repurposes them for obligation for evidence-building activities. Such creativity to support critical data infrastructure and evidence priorities for the American people should be encouraged. Importantly, these resources should supplement, rather than supplant, other shared capabilities and funding for the ICSP, CDO Council, EO Council, and other evidence-building activities.

**Recommendation 6.4. The OMB Director should prioritize additional resources for OMB staff responsible for coordinating implementation of the Evidence Act Title III regulations, Title II guidance, Title I implementation activities, and other evidence-building priorities in the current fiscal year and in the FY 2024 Budget request.**

Finally, the Committee recognizes that there are meaningful limitations in the capacity at OMB, which collaborates with agencies in approving and providing guidance on many aspects of implementing the Evidence Act, working with CDOs, EOs, and SOs. The Committee notes that notwithstanding the excellent contributions from the assigned staff for this Committee and the regulatory work under the Evidence Act, the regulations and guidance missed every statutory deadline set by Congress and agreed to by the Executive Branch when the law was enacted.

Underinvesting in these critical areas at OMB, given the Administration’s commitment to data and evidence priorities, would greatly limit successful implementation of the Evidence Act in coming years. Furthermore, additional resources to build staff capacity at OMB should be in balance with investments for statistical and nonstatistical agencies, so that the whole of government works better together. It would not be desirable, for example, to greatly fund the OMB staff tasked with delivering the regulations and guidance without providing sufficient resources to the departments and agencies required to implement them.

The initial guidance and regulatory development at OMB are paramount for implementing the Committee’s recommendations and vision and must be a priority for the OMB Director and the Office of Information and Regulatory Affairs Administrator in allocating resources and staff capacity in the years ahead. The success of the data service rests on the use of the regulatory authorities and forthcoming guidance that must be issued by OMB in a timely fashion with effective stakeholder engagement practices, and this can only be achieved with adequately prioritized staff capacity. For more information, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance.

## *State, Territorial, Local, and Tribal Governments*

**Recommendation 6.5. The OMB Director should propose in the FY 2024 Budget request a new block grant for state, territorial, local, and tribal funding to support cross-program data infrastructure improvements and data modernization.**

The lack of resources, capacity, and technical capability at the state, territorial, local, and tribal levels hinders the data collection process and limits data management, use, and analysis. The need for additional, sustained, and significant resources to incentivize data producers to integrate their data systems is a key challenge across all levels of government.

The Year 1 report described the two-way needs for state, territorial, local, and tribal data providers and users: “...despite the potential for evidence building, resources are not provided from the federal or state level down to the point-of-origin for data collection to support high-quality information flowing back up to federal, state, and local decision-makers.”

Furthermore, the Government Data for Evidence Building focus area stressed that: “...evidence building is enhanced when the data are available to staff at all levels of government. So, while the federal government must work with state and local governments to provide quality data, it also needs to ensure that state and local officials have access to the data once they go up to the federal databases.” The subcommittee also noted that data quality starts at the data source, which is often at the state, territorial, local, and tribal levels. To ensure that the federal government has the high-quality data it needs, federal agencies should support building capacity and capability among these entities. The Committee reaffirms and strengthens this statement: For there to be data and evidence, there need to be directly allocated resources for data management and training at the state and local levels.

During Year 2, the Committee discussed ongoing efforts aimed at providing federal funding to improve data infrastructures and analysis capabilities at the state, territorial, local, and tribal levels. Here are a few examples:

- Under the [CARES](#) Act, the National Center for Health Statistics National Vital Statistics System modernization initiative provides funding to all 57 jurisdictional partners that allows for more rapid receipt, coding, review, analysis, and release of data on births and deaths.
- Through the Department of Labor (DOL) Equity Data Partnerships program, states can voluntarily enter data-sharing agreements with DOL designed to improve the equitable delivery of unemployment compensation benefits through better data collection, measurement, and analysis.
- The OMB memorandum on “Advancing Effective Stewardship of Taxpayer Resources and Outcomes in the Implementation of the Infrastructure Investment and Jobs Act (IIJA)” ([M-12-22](#)) authorizes agencies to use infrastructure money for evidence-building activities and provides guidance for doing so. While this memo applies only to federal agencies, state and local governments could take a similar approach to using IIJA funding to improve data infrastructures and analytical tools in support of better policymaking.

The Committee acknowledges the success of existing programs and the potential of using sources like infrastructure funds to enhance state and local data systems; however, it also recognizes that funding substantial investments in human capital and technology across thousands of governments is a complicated task that no single existing mechanism, particularly a one-time funding mechanism, is likely to address.

Uptake of one-time funding for major investments in data infrastructure and modernization is also historically low because state, territorial, local, and tribal jurisdictions may be challenged to identify sustained investments for program operations, including those to support programs directed through cooperative federalism. Sustained funding mechanisms that support interoperable, coordinated, and integrated systems such as those used under the [OMB A-87](#) waivers for the Department of Health and Human Services and other agencies over the last decade, among other examples, offer compelling illustrations of the value of sustained investments for integrated capacity.

To provide a comprehensive and coordinated mechanism for funding data infrastructure and interoperability improvements, analytical tools, and capacity building for state, territorial, local, and tribal governments, the Committee strongly recommends the establishment of a new, sustained, and significant funding flow—like a data-related block grant and data infrastructure modernization fund.

Furthermore, the Committee explored how the agency assistance described in the [Secure Research Data Network Act](#) could be used as a model for administering these funds. Under this approach, agencies receive free assistance from technical, IT, and security experts to help connect their systems to the research network. The agencies are also reimbursed for their own staffs' time. In return, the agencies receiving assistance agree to implement the experts' recommendations and to provide access to relevant data.

The value of data modernization for state and local governments is significant. At the same time, the solution for one jurisdiction may not meet the needs of another jurisdiction. The federal government can do more to reduce unfunded mandates related to data standards and governance, accelerate the pace of data modernization, and realize improved data capabilities. In practice, these funding mechanisms must provide adequate support and assistance to state, territorial, local, and tribal governments that supplement, not replace, existing resource streams. Additional investment must yield resources that demonstrate value back to the American taxpayer.

## NSDS Resources to Meet Its Mission

### *Direct Spending Authority*

**Recommendation 6.6. NSDS core functions should be funded through direct spending authority.**

The NSDS is a public good, serving the broad evidence ecosystem. As such, its core services—coordination, communication, R&D, and data standardization—are critical infrastructure for ensuring that the NSDS is useful and used. It is vital that the NSDS receives the resources its needs to support its defined mission.

Although service providers, partners, and others in the evidence ecosystem may be transient to NSDS services (both as providers and users), NSDS operations require consistency. For more information on the role of service providers and partners, see Recommendations Part 3. NSDS Functions—Coordination (Support a High-Quality User Experience, Including Providing Technical Assistance and Recommendations Part 4. NSDS Organizational Structure and Governance—NSDS as a Government-Owned and Contractor-Operated Organization.

***Near-term strategy.*** To start building these core services, the Committee affirms the importance of the National Center for Science and Engineering Statistics (NCSES) FY 23 Budget request, which includes funding for America’s DataHub Consortium, the Standard Application Process, and early work on the NSDS. These resources would advance the establishment of the NSDS by amplifying the funding provided through the [CHIPS and Science Act](#). For more information on the NSDS funding structure, see Recommendations Part 4. NSDS Organizational Structure and Governance—America’s DataHub Consortium as the Foundation for the NSDS.

***Sustainable approach.*** Resources to establish the NSDS, however, are only the beginning. As stated in the Year 1 report, “for a data service to be successful, legislation may be needed to provide appropriate authority, scope, and funding for a National Secure Data Service.” In Year 2, the Committee echoes this message—louder and stronger: Resources for the NSDS are paramount to its success, along with the resources to support data infrastructure across the larger ecosystem for data users and producers that will interact with the NSDS.

Direct spending authority is a reliable, justifiable, and necessary funding mechanism for the NSDS because of the use of existing supports and the expectation for a consistent baseline of service delivery as a shared service for government agencies and the broader evidence ecosystem. The ability to offer new services and capabilities that supplement, rather than supplant or displace, existing resources, is a reasonable basis for new direct spending authority, and perhaps a reasonable basis for OMB to consider where other core statistical, evaluation, and data infrastructure activities should also be adjusted for direct spending authority across the federal budget.

These evidence-building activities are all expected, necessary, and essential for operating federal programs, monitoring program activities, and evaluating outcomes, and yet the Committee finds that underinvestment in basic data infrastructure persists. The NSDS core functions must be positioned to overcome these historic funding limitations.

The Committee also discussed how to scope and implement the direct funding authority for the NSDS. Key considerations are described below.

***Funding to support NSDS staffing, infrastructure, and contracted services.*** NCSES must have a specific allocation to support a talented and dedicated staff responsible for overseeing operations and ensuring that the data service meets its mission. NCSES should explore the appropriate staffing level, acknowledging that the size of the staff may need to expand over time as additional data service capabilities come online and user demand grows. In addition, direct funding should be allocated to developing technical infrastructure and tools and supporting core NSDS services provided by the primary contractor or other service providers. For more information, see Recommendations Part 3. NSDS Functions, Recommendations Part 4. NSDS Organizational Structure and Governance, and Recommendations Part 5. NSDS Technical Infrastructure and Tools.

***Cooperative stewardship model.*** A cooperative stewardship model, like the [one described by the National Academies of Sciences, Engineering, and Medicine](#), could increase interagency cooperation by providing incentives for diverse agencies across the federal government to leverage the services of the NSDS. Under this model, there would be three allocations—one for the NSDS (that is, for the NCSES Program Management Office), one for statistical and programmatic agencies already using NSDS services (or likely to do so), and one for open competition to give the potential for new agencies to get additional appropriations. The funds for the two agency allocations would be tagged so that they could only be used for NSDS activities. The NSDS should explore the cooperative stewardship model as part of its strategy for harnessing direct funding authority.

## *Mixed Funding Model*

**Recommendation 6.7.** As data service functions and user demand for its services grow over time, the NSDS should explore a mixed funding model that leverages sustainable and dynamic funding approaches, including budget requests through NCSES, existing and new federal grant programs, repurposed agency funds, federal-state partnerships, private-sector support, a shared services model, and user fees for select services.

Establishing the ADC as the foundation for the NSDS allows for flexible discretionary funding streams from a variety of sources. For more information, see Recommendations Part 3. NSDS Functions—Coordination (Project Sponsorship) and Recommendations Part 4. NSDS Organizational Structure and Governance—America’s DataHub Consortium as the Foundation for the NSDS.

Other funding models and approaches that the Committee explored, as well as key takeaways, are presented below.

***Funding through existing programs.*** Recognizing the multi-directional value stream that exists in federal-state partnerships, funding for NSDS services could also be channeled through specific programs, such as unemployment insurance and federal job training and education initiatives, since these programs would benefit from the coordination and capacity-building services of the NSDS, as illustrated in the Committee’s use cases. For more information, see Appendix B. ACDEB Use Cases and the use case reports in the Supplemental Information [posted](#) with this report.

***Public-private partnerships.*** While there may be a perception that resources for programs like Federal Statistical Research Data Centers (FSRDCs) come from the Census Bureau and other statistical agencies, in fact, partner institutions invest substantial resources into these efforts. While every FSRDC may like to provide support to less-resourced universities and non-academic researchers, the ability to do so depends greatly on the investment of partner institutions, which varies substantially across the network. While the FSRDC system does not have a mandate to provide services to “all comers,” for the NSDS to be successful in its mission, it must address issues like this.

***Shared services model.*** The National Center for Advancing Translational Science (NCATS) provides an example of how a shared services model can be used to support evidence building. NCATS offers free tools, resources, and training paired with a shared services model that leverages the federal government’s buying power to build a broadly accessible world-class resource. The NSDS should consider shared approaches for maximizing the services it provides.

***User fees for service.*** The Committee emphasizes that there is an appropriate role for user fees in the NSDS funding structure in line with [OMB Circular No. A-25](#)—not as a main funding stream to support core infrastructure and functions but to expand services by enabling additional programs or initiatives.

Flexibility to leverage user fees—and other mechanisms, like offsetting collections and gift authority—should be considered for the NSDS, along with authority to obligate and outlay those funds for approved projects and activities. Without this authority, such funds would be redirected to the U.S. Treasury and not reinvested in the data service. This echoes the Evidence Commission’s recommendation: “The NSDS should have the authority to collect and spend user fees, with sufficient flexibility to adjust rates based on changes in demand or other factors.”

There is a finite privacy budget related to each data asset, so there is a privacy “cost” to using and disseminating the data. In other words, making too many, too accurate disclosures of data increases the risk of re-identification. User fees should reflect this cost, tied to a risk-utility framework that reflects both the inherent privacy loss of using the data as well as the anticipated benefit of analysis. For more information on the risk-utility framework, see Recommendations Part 1. Forthcoming Evidence Act Regulations and Guidance—Expanding Secure Access to CIPSEA Data Assets (Risk and Utility Framework) and Recommendations Part 4. NSDS Organizational Structure and Governance—Holistic Governance Approach to Transparency and Accountability.

## 4. Appendices

### Appendix A: ACDEB Committee Charter, Membership, Process, and Meetings

This appendix provides an overview of the Committee’s charter, membership, process, tools, and meetings. Additional information related to the Committee can be found on [ACDEB’s website](#). Specific items are described below.

#### Charter

The Committee was established under the Evidence Act to review, analyze, and make recommendations to the Office of Management and Budget Director on how to promote the use of data for evidence building. ACDEB is chartered under the Federal Advisory Committee Act (FACA). As such, the [Committee’s charter](#) provides information on its authority, objectives and scope of activities, description of duties, agency or official to whom the Committee reports, support, estimated annual operating costs and staff years, Designated Federal Officer, estimated number and frequency of meetings, duration, termination, membership and designation, subcommittees, and recordkeeping.

#### Membership

The Committee comprises 26 members, representing diverse perspectives and a wealth of expertise from federal, state, and local governments as well as the private sector (including the privacy community).

A complete list of Committee members and their bios are available on the [ACDEB website](#).

## Process

The Committee’s process was structured to ensure the delivery of two annual reports, as mandated by its charter. The following section details this process during Year 1 and Year 2.

**Year 1.** ACDEB launched its work in October 2020 and spent the first 7 months sharing knowledge and experience relevant to fundamental evidence-building issues. To gather information, Committee members and outside experts shared presentations at public meetings, and the Committee solicited feedback from the public through a request for comment in the Federal Register.

From there, the Committee began to build on this knowledge base, breaking discussion into five focus areas and related subcommittees: (1) legislation and regulations; (2) governance, transparency, and accountability; (3) technical infrastructure; (4) government data for evidence building (with an emphasis on administrative data); and (5) other services and capacity-building opportunities. Drawing on the vast input and insights of Committee members, the focus areas were designed to address a wide range of opportunities as well obstacles for a National Secure Data Service and the evidence-building ecosystem of which it is a part.

In addition, the Committee established a coordinating committee to ensure consistency across focus areas and to minimize duplication of effort. This group comprises a cross-section of members from different focus areas with diverse expertise and experience.

Per FACA requirements, subcommittees presented their findings and recommendations to the full Committee at public meetings. Table A1 lists subcommittee co-chairs and members.

**Table A1. ACDEB Subcommittees**

Subcommittee	Members
Coordinating Committee	Emilda Rivers (chair), Laila Alequresh, Charles Cutshall, Nicholas Hart, Anna Hui, Julia Lane, Amy O’Hara (through May 2022), Mayank Varia (starting June 2022), Matthew Soldner, Kenneth Troske
Legislation and Regulations	Nicholas Hart (co-chair), Christine Heflin (co-chair), Gregory Fortelny, Ted Kaouk, Edward Kwartler, Christin Lotz, Todd Richardson, Mayank Varia
Governance, Transparency, and Accountability	Charles Cutshall (co-chair), Julia Lane (co-chair), Otis Brown, Shawn Davis, Gregory Fortelny, Edward Kwartler, Brian Moyer, Kimberly Murnieks, Christina Yancey
Technical Infrastructure	Amy O’Hara (co-chair), David Park (co-chair), Otis Brown, Leonard Burman, Barry Johnson, Ted Kaouk, Elisabeth Kovacs, Mayank Varia, Christina Yancey
Government Data for Evidence Building	Anna Hui (co-chair), Kenneth Troske (co-chair), Laila Alequresh, Richard Allen, Leonard Burman, Christine Heflin, Elisabeth Kovacs, Christin Lotz, Brian Moyer
Other Services and Capacity-Building Opportunities	Kimberly Murnieks (co-chair), Matthew Soldner (co-chair), Richard Allen, Leonard Burman, Shawn Davis, Barry Johnson, David Park, Todd Richardson

**Year 2.** The second-year timeline, deliverables, and workflows were designed to create building blocks for the Year 2 report, the Committee’s congressionally mandated final deliverable. Table A2 presents the Committee’s Year 2 roadmap.

**Table A2. ACDEB Year 2 Roadmap**

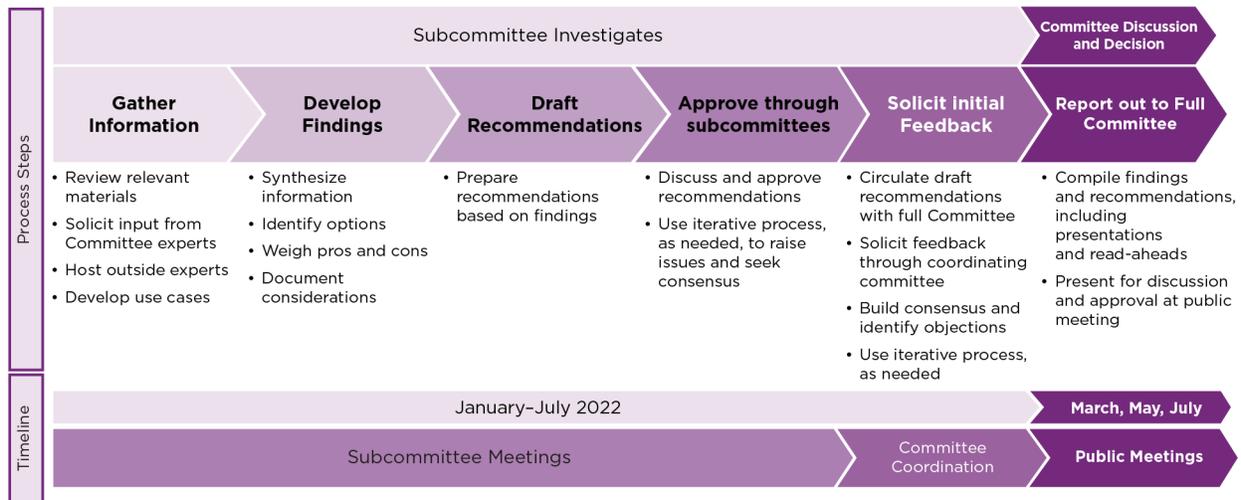
Date	Milestone	Notes
<b>October 29, 2021</b>	<b>Year 1 report delivered to OMB</b>	
November 4–18, 2021	Year 1 Reflection Sessions	<b>Topics:</b> What worked well? What didn’t?
<b>November 19, 2021</b>	<b>ACDEB Meeting 14</b>	<b>Topics:</b> Context and Year 2 Considerations from the Evidence Act (CIPSEA), NSDS Attributes and Functions, Year 2 Roadmap
Late Nov.–Early Dec. 2021	Subcommittee meetings restart	<b>Topics:</b> Focus Area Year 2 Roadmaps, explore links to the Evidence Act
Subcommittees gather information, develop findings, weigh options, and draft recommendations		
<b>January 21, 2022</b>	<b>ACDEB Meeting 15</b>	<b>Topic:</b> Introduction: Keeping Pace with the Evolving Federal Data Ecosystem, OMB and ICSP Response to ACDEB Year 1 Report, ICSP Update: Standard Application Process and Other Ongoing Projects, Structural Options for America’s DataHub Consortium, NSF Spotlight: America’s DataHub Consortium
Subcommittees gather information, develop findings, weigh options, draft recommendations, and prepare read-aheads and presentations for full Committee		
February 18, 2022	America’s DataHub Consortium Virtual Site Visit	Topic: Follow-up from January ACDEB meeting, open Q&A
March 3, 2022	Standard Application Process Virtual Site Visit	Topic: Follow-up from January ACDEB meeting, open Q&A
<b>March 18, 2022</b>	<b>ACDEB Meeting 16</b>	<b>Topics:</b> Subcommittee Reports, including findings and proposed recommendations
Subcommittees gather information, develop findings, weigh options, draft recommendations, prepare read-aheads and presentations for full Committee		
<b>May 20, 2022</b>	<b>ACDEB Meeting 17</b>	<b>Topics:</b> Recommendations Summary, Subcommittee Reports, including findings and proposed recommendations
Subcommittees gather information, develop findings, weigh options, draft recommendations, prepare read-aheads and presentations for full Committee		
<b>July 22, 2022</b>	<b>ACDEB Meeting 18</b>	<b>Topics:</b> Year 2 Report: Expectations, Timeline, and Deliverables; Facilitated Discussions: Resources and Funding; Governance, Transparency, and Accountability; Recommendations Summary
Finalize background materials, recommendations, findings, and supporting materials		
<b>September 23, 2022</b>	<b>ACDEB Meeting 19</b>	<b>Topics:</b> Year 2 Report: Overview of Report Framework and Major Changes Since July Meeting, Outstanding Report Items and Options, Committee Reflections and the Next Chapter for Advancing the Data Evolution, Outreach and Stakeholder Engagement
October 14, 2022	Deliver Year 2 report to OMB	
<b>October 21, 2022</b>	<b>ACDEB Meeting 20 (in person)</b>	<b>Topics:</b> Committee close out and final thoughts
<b>October 22, 2022</b>	<b>Committee charter expires</b>	

- ACDEB Advisory Committee on Data for Evidence Building
- CIPSEA Confidential Information Protection and Statistical Efficiency
- ICSP Interagency Council on Statistical Policy
- NSDS National Secure Data Service
- NSF National Science Foundation
- OMB Office of Management and Budget

The subcommittees carried the bulk of the workload during Year 2, organized in several phases—gather information, develop findings, draft recommendations, approve through subcommittees, solicit initial feedback from full Committee, and report out at the public meeting—and compressed into the period from January to July 2022. As part of the information-gathering phase, subcommittees developed use cases, engaged in iterative conversations with the Office of Management and Budget (OMB) and the Interagency Council on Statistical Policy (ICSP), and hosted meetings with outside experts. In addition, the Committee sponsored field trips, open to all members, on topics of broad interest. Throughout this process, the coordinating committee facilitated information sharing across the subcommittees, including providing initial input, soliciting feedback, building consensus, and identifying objections using an iterative process, as necessary.

Table A3 shows the focus area workflow, including highlighting the role of the coordinating committee. For more information, see the other appendixes and the Supplemental Information [posted](#) with this report.

**Diagram A1. Focus Area Workflow**



## Tools

The Committee developed and leveraged several tools to help inform its work, including a [resource library](#), a [project inventory](#), a use case proposal template, and a crosswalk between the Year 1 report and Title III of the Evidence Act (or CIPSEA 2018). For more information on the use cases, see Appendix B and the Supplemental Information—Use Case Report [posted](#) with this report. For details on the crosswalk, see the “[Context and Year 2 Considerations from the Evidence Act \(CIPSEA\)](#)” presentation from the November 2022 ACDEB meeting.

## Meetings

The content and structure of the Committee’s public meetings mirrored its phased process. For Year 1, the first seven meetings focused on information gathering. The remaining meetings centered on subcommittee reports and Committee discussions around process and possible recommendations. For Year 2, the full Committee met every other month. The first two meetings used the Year 1 report as a springboard for next steps—diving deeper into NSDS attributes and functions, highlighting the Year 2 roadmap, kicking off the iterative conversation with OMB/ICSP, and exploring the Standard Application Process and America’s DataHub Consortium as models for the NSDS. The remaining meetings in Year 2 focused on subcommittee reports and gathering consensus around Committee recommendations and final report content.

Meeting agendas, presentations, and minutes are available on the [ACDEB website](#).

## Appendix B. ACDEB Use Cases

During its second year, the Committee explored several use cases as a mechanism for gathering evidence and developing findings to support its recommendations. As part of this process, members considered each of the items below and addressed them, as applicable.

### Focus

- Review current and evolving approaches to accessing, linking, and analyzing data across the federal, state, and local levels with a consideration of how decisionmaking could be enhanced and facilitated.
- Investigate improvements for the current evidence-building ecosystem and weigh the possibilities of a National Secure Data Service.

### Rationale

- Evidence Act with an emphasis on Title III (that is, the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) 2018)
- Evidence Commission recommendations
- ACDEB Year 1 report (full Committee and subcommittee recommendations)
- Interagency Council on Statistical Policy (ICSP) and Office of Management and Budget (OMB) workstreams

### Key Points

- The value of data access, linking, and analysis for evidence building for data providers and users
- Potential data sources: Federal, state, local, and private sources; both official statistical products and administrative data
- Barriers, challenges, and gaps; for example:
  - Legal and regulatory barriers;
  - Cultural resistance;
  - Differences in metadata, data quality, and systems interoperability; and
  - Resource and capacity issues.
- Lessons learned, possible solutions, and opportunities; for example:
  - Coordination within and across levels of government and with the private sector;
  - Data standards, consistency, and interoperability; data quality improvements; and data discoverability and transparency;
  - Role of the academic community, communities of practice, training, and resource sharing; and
  - Technologies, tools, and advanced analytical methods
- Privacy and confidentiality: Privacy-quality tradeoff and privacy protections needed to comply with legal and ethical requirements

Table B1 describes each use case. For the full use case reports, see the Supplemental Information [posted](#) with this report.

**Table B1. ACDEB Use Cases**

Project Title and Purpose	Models of focus	Champions	Originating Subcommittee
<p><b>Using Administrative Data to Track Project Impact</b> Purpose: To explore the use of administrative data to track project outcomes, especially at the local level.</p>	<ul style="list-style-type: none"> <li>Implementation of American Rescue Program and Infrastructure Investment and Jobs Act</li> </ul>	<ul style="list-style-type: none"> <li>Christine Heflin</li> </ul>	Legislation and Regulations
<p><b>Education and Workforce</b> Purpose: To review current and evolving approaches to accessing, linking, and analyzing education and workforce data across federal, state, and local levels with a consideration of how decisionmaking could be enhanced and facilitated. This includes investigating improvements for the current evidence-building ecosystem and weighing the possibilities for the NSDS.</p>	<ul style="list-style-type: none"> <li>Department of Education Ability to Benefit program</li> <li>The Midwest Collaborative</li> </ul>	<ul style="list-style-type: none"> <li>Shawn Davis</li> <li>Gregory Fortelny</li> </ul>	Governance, Transparency, and Accountability
<p><b>Health</b> Purpose: To review current and evolving approaches to accessing, linking, and analyzing health data across federal, state, and local levels with a consideration of how decisionmaking could be enhanced and facilitated. This includes investigating improvements for the current evidence-building ecosystem and weighing the possibilities for the NSDS.</p>	<ul style="list-style-type: none"> <li>National Center for Health Statistics National Vital Statistics System modernization efforts</li> </ul>	<ul style="list-style-type: none"> <li>Brian Moyer</li> <li>Kimberly Murnieks</li> </ul>	Governance, Transparency, and Accountability
<p><b>Labor Market Activity</b> Purpose: To review current and evolving approaches to accessing, linking, and analyzing labor market activity data with an emphasis on (1) improving local labor market statistics, (2) leveraging unemployment insurance data for better evaluation/research/continuous improvement, and (3) producing new national statistics. This includes investigating improvements for the current evidence-building ecosystem and weighing the possibilities for the NSDS.</p>	<ul style="list-style-type: none"> <li>Department of Labor Unemployment Insurance (UI) Equity Data Partnerships</li> <li>Regional state collaboratives</li> </ul>	<ul style="list-style-type: none"> <li>Julia Lane</li> <li>Christina Yancey</li> </ul>	Governance, Transparency, and Accountability
<p><b>Health Quality and Human Health</b> Purpose: To explore approaches for better accessing, linking, and analyzing data for environmental condition and human health.</p>	<ul style="list-style-type: none"> <li>Per- and polyfluoroalkyl substances contamination in drinking water</li> </ul>	<ul style="list-style-type: none"> <li>Richard Allen</li> </ul>	Government Data for Evidence Building
<p><b>Data Inventories and Metadata</b> Purpose: To review current and evolving approaches to creating data inventories and managing metadata, with a consideration of how such tools could enhance and facilitate evidence-based decisionmaking at federal, state, and local levels.</p>	<ul style="list-style-type: none"> <li>U.S. Chamber of Commerce Foundation Jobs and Employment Data Exchange</li> </ul>	<ul style="list-style-type: none"> <li>Elisabeth Kovacs</li> </ul>	Government Data for Evidence Building

## Appendix C. ACDEB Virtual Site Visits

As part of the information-gathering process during Year 2, ACDEB sponsored two Committee-wide virtual site visits—one focused on America’s DataHub Consortium (ADC) and another on the Standard Application Process (SAP). Members raised many questions on these topics during and after the January 2022 ACDEB meeting. These field trips provided another opportunity for members to engage with relevant outside experts.

Table C1 provides an overview of each site visit. The Committee would like to thank all speakers and support staff who made these events a reality. The information shared during these visits does not reflect the views of the full Committee. For summaries of the site visits, see the Supplemental Information [posted](#) with this report.

**Table C1. Virtual Site Visits Overview**

Virtual Site Visit	Host(s), Speaker(s), and Support Staff
<b>America’s DataHub Consortium</b> February 18, 2022	Speakers: Vipin Arora (NCSES), Keith Boyea (NSF), Dolly Peltó (ATI), John Finamore (NCSES)
<b>Standard Application Process</b> March 3, 2022	Speakers: Vipin Arora (NCSES), John Eltinge, John Finamore (NCSES), Alex Marten (EPA), Spiro Stefanou (USDA-ERS)

ATI      Advanced Technology International  
 EPA      Environmental Protection Agency  
 ERS      Economic Research Service  
 NCSES   National Center for Science and Engineering Statistics  
 NSF      National Science Foundation  
 USDA    United States Department of Agriculture

## Appendix D. OMB/ICSP Workstreams

At the January 2022 ACDEB meeting, OMB and the ICSP provided the overview for iterative engagement with the Committee. These “workstreams” offered the Committee the opportunity (1) to provide feedback on federal statistical system initiatives that were already in progress and (2) to use these existing efforts to inform the Committee’s work. These two pieces support the Committee’s charge to advise OMB on CIPSEA (Title III of the Evidence Act)—both on the “here and now” and on the target vision for the National Secure Data Service.

On any given workstream, there may have been a series of discussions with the primary subcommittee (or members thereof) and other members, as appropriate. As each conversation evolved, members provided preliminary input through discussion, answers to targeted questions, written comments, or other informal mechanisms. This initial input filtered through to the findings and recommendations presented in this report.

Table D1 presents the OMB/ICSP workstreams, provides a brief description of each topic, lists the “assigned” ACDEB subcommittee(s), and notes the ICSP leaders for each workstream. The Committee would like to thank OMB and the ICSP for this engagement. The information shared during these sessions does not reflect the views of the full Committee. For summaries of the OMB/ICSP sessions with ACDEB’s subcommittees, including key takeaways, see the Supplemental Information [posted](#) with this report.

**Table D1. OMB/ICSP Workstreams**

Workstream	Description	ACDEB subcommittee	OMB/ICSP leaders
Access and Confidentiality Initiative: Regulation	<ul style="list-style-type: none"> <li>Elicit feedback on goals, principles, and frameworks.</li> <li>Build shared understanding of challenges.</li> <li>Identify innovative ways to meet the goals and overcome challenges.</li> <li>Includes data sensitivity levels that are also a part of the new Zero Trust policy development in which the CDO Council is engaged.</li> </ul>	Legislation and Regulations	Lead Champion: Spiro Stefanou Supporting: Shelly Martinez
Access and Confidentiality Initiative: Methods Coordination	<ul style="list-style-type: none"> <li>Gather input and feedback on priorities to build out Data Protection Toolkit.</li> <li>Validate what ACDEB is looking for in case study narratives/pilot goals that cover value-driven projects and privacy-preserving technologies.</li> <li>Provide feedback and validation on methods coordination activities and how an NSDS could address those activities.</li> </ul>	Technical Infrastructure	Lead Champion: Barry Johnson Supporting: Michael Hawes
Stakeholder Engagement: Access and Confidentiality and Standard Application Process Initiatives	<ul style="list-style-type: none"> <li>Solicit advice on the essential elements of a successful engagement plan, especially how to effectively engage nonfederal users.</li> <li>Ask for feedback on key stakeholder messaging.</li> <li>Inform ACDEB on stakeholder engagement work and gather consensus on messaging that highlights data access, linkage, and data protection methods.</li> <li>Share best practices on agency collaboration.</li> <li>Gather suggestions for continued synchronized messaging.</li> </ul>	Other Services and Capacity-Building Opportunities Supporting: Government Data for Evidence Building	Lead Champion: Bill Beach Supporting (Access and Confidentiality): Guadalupe Cerritos Supporting (SAP): Vipin Arora
Standard Application Process Initiative: Technical Development	<ul style="list-style-type: none"> <li>Gather input on the technical requirements to develop and implement the process required under Section 3582.</li> <li>Discuss initial technical requirements for the SAP, including the vision for metadata and how tiered access requirements relate to SAP implementation.</li> </ul>	Technical Infrastructure	Lead Champion: Barry Johnson Supporting: Shelly Martinez
Standard Application Process Initiative: Governance	<ul style="list-style-type: none"> <li>Seek advice on the governance structure that the ICSP is looking to put into place prior to the launch of Phase 2 implementation, consistent with the draft policy proposal (FRN).</li> </ul>	Governance, Transparency, and Accountability	Lead Champion: Brian Moyer Supporting: Alex Marten, Spiro Stefanou
Standard Application Process: Technical Assistance	<ul style="list-style-type: none"> <li>Gather feedback on how to incorporate more user support features in future releases.</li> </ul>	Other Services and Capacity-Building Opportunities	Lead Champion: Barry Johnson Supporting: Shelly Martinez
Other Items	<p>Other items upon which advice could be useful and for which OMB/ICSP can provide topics/questions if ACDEB is interested in engaging:</p> <ul style="list-style-type: none"> <li>Responsibilities of Statistical Agencies (aka Trust) regulation</li> <li>Chief Statistician Priorities</li> <li>FY 23 Appropriations</li> <li>Standards Setting</li> </ul>	Legislation and Regulations Legislation and Regulations Government Data for Evidence Building Governance, Transparency, and Accountability	Supporting: Dominic Mancini, Shelly Martinez, Robert Sivinski

## Appendix E. ACDEB Subcommittee Guest Speakers

As part of the information-gathering process during Year 2, ACDEB subcommittees hosted targeted discussions with outside experts. To cross-pollinate ideas from the focus areas, other Committee members were invited to attend these sessions and ask questions from the perspectives of their subcommittees and areas of expertise.

Table E1 provides an overview of each outside expert meeting. The Committee would like to thank all speakers and supporting staff who made these sessions a reality. The information shared during these discussions does not reflect the views of the full Committee. For the related meeting summaries, see the Supplemental Information [posted](#) with this report.

**Table E1. Outside Expert Meetings Overview**

Organization(s), Topic, and Date	Host Subcommittee	Speaker(s), Planning, and Support
<b>Data Quality Campaign: Communication strategies</b> February 24, 2022	Other Services and Capacity-Building Opportunities	Rachel Anderson, Jenn Bell-Ellwanger
<b>Midwest Collaborative (MWC), National Association of State Workforce Agencies (NASWA), and Workforce Information Advisory Council (WIAC): Governance insights</b> March 4, 2022	Governance, Transparency, and Accountability	George Putnam (MWC and Illinois Department of Employment Security), Yvette Chocolaad (NASWA), Lesley Hirsch (WIAC and New Jersey Department of Labor and Workforce Development)
<b>Results for America: Communication strategies</b> March 10, 2022	Other Services and Capacity-Building Opportunities	Cheryl Burnett, Zachary Coile, Nichole Dunn
<b>Urban Institute: Synthetic Data and Validation Servers</b> April 8, 2022	Technical Infrastructure	Leonard Burman (ACDEB and Urban Institute), Graham MacDonald
<b>Datavant: COVID-19 Research Database</b> April 8, 2022	Technical Infrastructure	Claire Cravero, Jake Plummer
<b>Jobs and Employment Data Exchange: Data dictionary</b> April 20, 2022	Government Data for Evidence Building	Kenneth Poole (Center for Regional Economic Competitiveness)
<b>Privacy threats and re-identification risks</b> April 21, 2022	Technical Infrastructure	Claire Bowen (Urban Institute)
<b>Federal Statistical Research Data Centers Panel: Decisionmaking, infrastructure, and technical assistance</b> April 21, 2022	Other Services and Capacity-Building Opportunities	Mary Campbell (Texas A&M), Barbara Downs (Census Bureau), Cathy Fitch (University of Minnesota), Maggie Levenstein (University of Michigan), Amy O'Hara (ACDEB and Georgetown University)
<b>State Wage Interchange System</b> May 4, 2022	Government Data for Evidence Building	Greg Wilson (Department of Labor), John (Jay) LeMaster (Department of Education)
<b>Inter-university Consortium for Political and Social Research: Technical assistance</b> May 5, 2022	Other Services and Capacity-Building Opportunities	Maggie Levenstein
<b>Privacy-preserving solutions for the future and risk evaluations</b> May 6, 2022	Technical Infrastructure	Wade Shen (Actuate Innovation)
<b>Opportunity Insights</b> May 6, 2022	Technical Infrastructure	John Friedman (Brown University)
<b>National Institutes of Health (NIH) Library Data Services: Technical assistance</b> May 19, 2022	Other Services and Capacity-Building Opportunities	John Doyle
<b>NIH National Center for Advancing Translational Sciences National COVID Cohort Collaborative</b> June 3, 2022	Technical Infrastructure	Kenneth Gersing, Sam Michael, Leonie Misquitta

# References

- “Leveraging Federal Statistics to Strengthen Evidence-Based Decision-Making,” *Fiscal Year 2023 President’s Budget, Analytical Perspectives of the Budget, Chapter 15*, [https://www.whitehouse.gov/wp-content/uploads/2022/03/ap\\_15\\_statistics\\_fy2023.pdf](https://www.whitehouse.gov/wp-content/uploads/2022/03/ap_15_statistics_fy2023.pdf).
- “Presidential Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking,” January 27, 2021, <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/>.
- “U.S. and U.K. Launch Innovation Prize Challenges in Privacy-Enhancing Technologies to Tackle Financial Crime and Public Health Emergencies,” July 20, 2022, <https://www.whitehouse.gov/ostp/news-updates/2022/07/20/u-s-and-u-k-launch-innovation-prize-challenges-in-privacy-enhancing-technologies-to-tackle-financial-crime-and-public-health-emergencies/>.
- Access 4 Learning Community, “Student Data Privacy Consortium,” accessed October 6, 2022, <https://privacy.a4l.org/>.
- Actuate Innovation, “DataSafes,” accessed October 6, 2022, <https://actuateinnovation.org/programs/datasafes>.
- Advanced Education Research and Development Fund, accessed October 6, 2022, <https://aerdf.org/>.
- Advisory Committee on Data for Evidence Building, accessed October 6, 2022, [www.bea.gov/evidence](http://www.bea.gov/evidence).
- Advisory Committee on Data for Evidence Building, “Advisory Committee on Data for Evidence Building: Year 1 Report,” *Bureau of Economic Analysis*, October 29, 2021. <https://www.bea.gov/system/files/2021-10/acdeb-year-1-report.pdf>.
- Advisory Committee on Evidence Building, “ACDEB Project Inventory,” *Bureau of Economic Analysis*, May 3, 2022, ACDEB Project Inventory, <https://www.bea.gov/system/files/2021-10/acdeb-project-inventory.xlsx>.
- Advisory Committee on Data for Evidence Building, charter, 2021, <https://www.bea.gov/system/files/2021-01/84587-DOC-2021-Charter-ACDEB-1.13.2021.pdf>.

- Alliance for Innovative Regulation, “The First US Regulatory TechSprint: Financial Crime and Human Trafficking,” accessed October 6, 2022, <https://regulationinnovation.org/techsprints/dc-regulatory-techsprint-on-aml-2/>.
- America’s DataHub Consortium, accessed October 6, 2022, <https://americasdatahub.org>.
- Apple, “Apple Differential Privacy Technical Overview,” accessed October 6, 2022, [https://www.apple.com/privacy/docs/Differential\\_Privacy\\_Overview.pdf](https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf).
- Archer, David, Amy O’Hara, Rawane Issa, and Stephanie Straus, “Sharing Sensitive Department of Education Across Organizational Boundaries Using Secure Multiparty Computation,” May 2021, <https://bit.ly/SMCatNCES>.
- Barnum, Sean and Michael Gegick, “Knowledge Topics: Least Privilege,” *Cybersecurity and Infrastructure Security Agency*, May 10, 2013, <https://www.cisa.gov/uscert/bsi/articles/knowledge/principles/least-privilege>.
- Boston Women’s Workforce Council, “Data: Gender and Racial Wage Gaps in Boston by the Numbers,” <https://thebwwc.org/wage-gap-studies>.
- Bureau of Labor Statistics, “Injuries, Illnesses, and Fatalities: Acquiring the Offsite CFOI Micro Fatality Research File,” September 9, 2020, [https://www.bls.gov/iif/cfoi\\_offsite.htm](https://www.bls.gov/iif/cfoi_offsite.htm).
- Bureau of Labor Statistics, “National Longitudinal Surveys: The Geocode Application Document,” June 6, 2022, <https://www.bls.gov/nls/geocodeapp.htm>.
- Census Bureau, “2020 Decennial Census: Processing the Count: Disclosure Avoidance Modernization,” accessed October 6, 2022, <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html>.
- Census Bureau, “Federal Statistical Research Data Centers,” August 9, 2022, <https://www.census.gov/about/adrm/fsrdc.html>.
- Census Bureau, “Longitudinal Employer-Household Dynamics,” accessed October 6, 2022, <https://lehd.ces.census.gov/>.
- Census Bureau, “Understanding Differential Privacy,” accessed October 6, 2022, <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/differential-privacy.html>.
- Centers for Medicare and Medicaid Services, “Alliance to Modernize Healthcare FFRDC,” December 1, 2021, <https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/CAMH>.
- Chief Data Officer Council, accessed October 6, 2022, <https://www.cdo.gov/>.
- CHIPS and Science Act, Pub. L. No. 117-167, 2022, <https://www.congress.gov/bill/117th-congress/house-bill/4346/text>.

- Coleridge Initiative, accessed October 6, 2022, <https://coleridgeinitiative.org/>.
- Coleridge Initiative, “Administrative Data Research Facility,” accessed October 6, 2022, <https://coleridgeinitiative.org/adrf/>.
- Coleridge Initiative, “Applied Data Analytics Training,” accessed October 6, 2022, <https://coleridgeinitiative.org/training/>.
- Coleridge Initiative, “Democratizing our Data: A Challenge to Invest in Data and Evidence-based Policy,” accessed October 6, 2022, <https://coleridgeinitiative.org/democratizing-our-data-challenge/>.
- Coleridge Initiative, “Midwest Collaborative: Data for Evidence-Based Policy (Spring Convening),” accessed October 6, 2022, <https://coleridgeinitiative.org/workshops/workshop-mar2020/>.
- Coleridge Initiative, “Multi-State Postsecondary Dashboard, Coleridge Initiative,” accessed October 6, 2022, <https://coleridgeinitiative.org/projects-and-research/multi-state-post-secondary-dashboard>.
- Coleridge Initiative, “Show US the Data,” accessed October 6, 2022, <https://coleridgeinitiative.org/projects-and-research/rich-context/>.
- Coleridge Initiative, “Unemployment to Reemployment Portal,” accessed October 6, 2022, <https://coleridgeinitiative.org/projects-and-research/unemployment-to-reemployment-portal/>.
- Commission on Evidence-Based Policymaking, “The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking,” Bipartisan Policy Center, September 2017, <https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Full-Report-The-Promise-of-Evidence-Based-Policymaking-Report-of-the-Commission-on-Evidence-based-Policymaking.pdf>.
- Coronavirus Aid, Relief, and Economic Security Act, Pub. L. No. 116-136, 2020, <https://www.congress.gov/bill/116th-congress/house-bill/748/text>.
- COVID-19 Research Database, accessed October 6, 2022, <https://covid19researchdatabase.org>.
- Cunningham, Jessica, Anna Hui, Julia Lane, and George Putnam, “A Value-Driven Approach to Building Data Infrastructures: The Example of the Midwest Collaborative,” *Harvard Data Science Review*, January 27, 2022, <https://hdsr.mitpress.mit.edu/pub/mfhfwpqxq/release/2>.
- Data.gov, accessed October 6, 2022), <https://data.gov>.
- Defense Advanced Research Projects Agency, “Brandeis,” accessed October 6, 2022, <https://www.darpa.mil/program/brandeis>.

- Defense Advanced Research Projects Agency, “Data Protection in Virtual Environments (DPRIVE),” accessed October 6, 2022, <https://www.darpa.mil/program/data-protection-in-virtual-environments>.
- Defense Advanced Research Projects Agency, “PROgramming Computation on EncryptED Data (PROCEED),” accessed October 6, 2022, <https://www.darpa.mil/program/programming-computation-on-encrypted-data>.
- Defense Advanced Research Projects Agency, “SAFEWARE,” accessed October 6, 2022, <https://www.darpa.mil/program/safeware>.
- Defense Advanced Research Projects Agency, “Securing Information for Encrypted Verification and Evaluation,” Defense Advanced Research Projects Agency, accessed October 6, 2022, <https://www.darpa.mil/program/securing-information-for-encrypted-verification-and-evaluation>.
- Department of Labor, “Department of Labor Equity Action Plan,” accessed October 6, 2022, <https://www.dol.gov/general/equity-action-plan/plan>.
- Department of Labor, Employment and Training Administration, “Wage Interchange Systems,” accessed October 6, 2022, <https://www.dol.gov/agencies/eta/performance/swis>.
- Department of the Treasury, “General Explanations of the Administration’s Fiscal year 2023 Revenue Proposals,” pp. 78-79, March 2022, <https://home.treasury.gov/system/files/131/General-Explanations-FY2023.pdf>.
- Evaluation Officer Council, September 1, 2021, <https://www.evaluation.gov>.
- Evidence-Based Policymaking Commission Act of 2016, Pub. L. No. 114-140, 2016, <https://www.congress.gov/bill/114th-congress/house-bill/1831/text>.
- Fatherree, Kira and Nicholas R. Hart, “Funding the Evidence Act: Options for Allocating Resources to Meet Emerging Data and Evidence Needs in the Federal Government,” *Data Foundation*, November 2019, <https://www.datafoundation.org/funding-the-evidence-act-paper-2019>.
- Federal Committee on Statistical Methodology, “A Framework for Data Quality,” September 2020, [https://nces.ed.gov/FCSM/pdf/FCSM.20.04\\_A\\_Framework\\_for\\_Data\\_Quality.pdf](https://nces.ed.gov/FCSM/pdf/FCSM.20.04_A_Framework_for_Data_Quality.pdf).
- Federal Committee on Statistical Methodology, “Data Protection Toolkit Beta,” September 22, 2020, <https://nces.ed.gov/fcsm/dpt>.
- Federal Privacy Council, accessed October 6, 2022, <https://www.fpc.gov/>.
- FedRAMP, “FedRAMP Marketplace,” accessed October 6, 2022, <https://marketplace.fedramp.gov/#!/products>.
- Foundations for Evidence-Based Policymaking Act of 2018, Pub. L. No. 115–435, 2019, <https://www.congress.gov/bill/115th-congress/house-bill/4174/text>.

- GO FAIR, “Fair Principles,” accessed October 6, 2022, <https://www.go-fair.org/fair-principles/>.
- Grumbling, Emily, Lisa Van Pay, and Morgan Livingston, “Options for Governance, Administration, and Ownership of a National AI Research Resource,” STPI, Editor, 2022.
- Hart, Nick, David Archer, and Erin Dalton, “Privacy Preserved Data Sharing for Evidence-Based Policy Decisions: A Demonstration Project Using Human Services Administrative Records for Evidence-Building Activities,” March 2019, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3808054](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3808054).
- Hart, Nick and Nancy Potok, “Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building,” *Data Foundation*, July 2020, <https://www.datafoundation.org/modernizing-us-data-infrastructure-2020>.
- Hart, Nick and Steve Mumford, “Federal Evaluation Officials 2021: 2021 Survey Results on Agencies’ Emerging Evaluation Capacity to Implement the Evidence Act,” *Data Foundation*, December 2021, <https://www.datafoundation.org/survey-of-federal-evaluation-officials-2021>.
- Hart, Nick, Tracy Jones, Jeff Lawton, Leigh Sheldon, Joe Willey, “CDO Insights: 2021 Survey Results on the Maturation of Data Governance in U.S. Federal Agencies,” *Data Foundation*, September 2022, <https://www.datafoundation.org/cdo-insights-report-2021>.
- High Conservation Value Network, “HCV Approach: Our Methodology Protects High Conservation Values from the Impacts of Land-Use Change,” accessed October 6, 2022, <https://www.hcvnetwork.org/hcv-approach>.
- Intelligence Research Projects Activity, “Security and Privacy Assurance Research (SPAR),” accessed October 6, 2022, <https://www.iarpa.gov/research-programs/spar>.
- Inter-university Consortium of Political and Science Research, accessed October 6, 2022, <https://www.icpsr.umich.edu/web/pages/>.
- JASON, “Secure Computation for Business Data,” November 2020, <https://irp.fas.org/agency/dod/jason/secure-comp.pdf>.
- Kaggle, “Coleridge Initiative - Show US the Data,” June 2021, <https://www.kaggle.com/c/coleridgeinitiative-show-us-the-data/>.
- Kurakin, Alexey, “Applying Differential Privacy to Large Scale Image Classification,” *Google AI Blog*, February 1, 2022, <https://ai.googleblog.com/2022/02/applying-differential-privacy-to-large.html>.
- Miklau, George, Michael Hay, Ashwin Machanavajihala, Amritha Pai, and William Sexton, “Negotiating Privacy-Utility Trade-Offs under Differential Privacy,” June 23, 2022, [https://www.usenix.org/system/files/pepr22\\_slides\\_miklau.pdf](https://www.usenix.org/system/files/pepr22_slides_miklau.pdf).

- Mirel, Lisa B., Dean M. Resnick, Jonathan Aram, and Christine S. Cox, “A Methodological Assessment of Privacy Preserving Record Linkage Using Survey and Administrative Data,” *Statistical Journal of the International Association for Official Statistics*, vol. 38, no. 2, pp. 413-421, 2022, <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji210891>.
- Modernizing Government Technology Act of 2017, H.R. 2227, 115th Cong. (2017), <https://www.congress.gov/bill/115th-congress/house-bill/2227>.
- National Academies of Sciences, Engineering, and Medicine, “Automated Research Workflows For Accelerated Discovery: Closing the Knowledge Discovery Loop,” 2022, Washington, D.C.: The National Academies Press, <https://nap.nationalacademies.org/catalog/26532/automated-research-workflows-for-accelerated-discovery-closing-the-knowledge-discovery>.
- National Academies of Sciences, Engineering, and Medicine, “Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies,” 2022, Washington, D.C.: National Academies Press, <https://nap.nationalacademies.org/read/26360/chapter/1>.
- National Academy of Public Administration, “Summary of Agile Government Principles,” April 6, 2020, <https://napawash.org/grand-challenges-blog/summary-agile-government-principles>.
- National Artificial Intelligence Research Resource Task Force, accessed October 6, 2022, <https://www.ai.gov/nairrtf/>.
- National Center for Health Statistics, “Modernizing the National Vital Statistics System,” October 1, 2020, <https://www.cdc.gov/nchs/nvss/modernization.htm>.
- National Center for Health Statistics, “National Vital Statistics System,” September 30, 2022, <https://www.cdc.gov/nchs/nvss/index.htm>.
- National Center for Health Statistics, “Research Data Center,” June 21, 2022, <https://www.cdc.gov/rdc/index.htm>.
- National Institutes of Health Library, “Data Services,” accessed October 6, 2022, <https://www.nihlibrary.nih.gov/services/data>.
- National Institutes of Health, National Center for Advancing Translational Science, “National COVID Cohort Collaborative,” September 26, 2022, <https://ncats.nih.gov/n3c>.
- National Opinion Research Center, *NORC at the University of Chicago*, October 6, 2022, <https://www.norc.org>.
- National Research Council, “Cooperative Stewardship: Managing the Nation’s Multidisciplinary User Facilities for Research with Synchrotron Radiation, Neutrons, and High Magnetic Fields,” Washington, D.C.: The National Academies Press, 1999, <https://nap.nationalacademies.org/catalog/9705/cooperative-stewardship-managing-the-nations-multidisciplinary-user-facilities-for-research>.

- Office of Management and Budget, “Advancing Effective Stewardship of Taxpayer Resources and Outcomes in the Implementation of the Infrastructure Investment and Jobs Act (IIJA),” (M-12-22), April 29, 2022, <https://www.whitehouse.gov/wp-content/uploads/2022/04/M-22-12.pdf>.
- Office of Management and Budget, Circular No. A-11, 2016, [https://obamawhitehouse.archives.gov/omb/circulars\\_all\\_current\\_year\\_all\\_toc](https://obamawhitehouse.archives.gov/omb/circulars_all_current_year_all_toc).
- Office of Management and Budget, Circular No. A-25, [https://obamawhitehouse.archives.gov/omb/circulars\\_a025/](https://obamawhitehouse.archives.gov/omb/circulars_a025/).
- Office of Management and Budget, Circular No. A-87, 2004, [https://obamawhitehouse.archives.gov/omb/circulars\\_a087\\_2004](https://obamawhitehouse.archives.gov/omb/circulars_a087_2004).
- Office of Management and Budget, “Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance” (memorandum M-19-23), July 10, 2019, <https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf>.
- Office of Management and Budget, Office of Information and Regulatory Affairs, “The Interagency Council on Statistical Policy’s Recommendation for a Standard Application Process for Requesting Access to Certain Confidential Data,” January 14, 2022, <https://www.federalregister.gov/documents/2022/01/14/2022-00620/the-interagency-council-on-statistical-policys-recommendation-for-a-standard-application-process-sap>.
- Office of Management and Budget, Office and Science Technology Policy, “Request for Information on Advancing Privacy-Enhancing Technologies,” *Federal Register*, June 9, 2022, <https://www.federalregister.gov/documents/2022/06/09/2022-12432/request-for-information-on-advancing-privacy-enhancing-technologies>.
- Opportunity Insights, September 12, 2022, <https://opportunityinsights.org/>.
- Paperwork Reduction Act, Pub. L. No. 104-13, 1995, <https://www.reginfo.gov/public/reginfo/pra.pdf>.
- Pediatric Data Commons, accessed October 6, 2022, <https://commons.cri.uchicago.edu/>.
- Potok, Nancy and Nick Hart, “A Blueprint for Implementing the National Secure Data Service: Initial Governance and Administrative Priorities for the National Science Foundation,” *Data Foundation*, June 2022, <https://www.datafoundation.org/cover-page-a-blueprint-for-implementing-the-national-secure-data-service>.
- Privacy Act, 5 USC § 552a(b)(4), 1974, <https://www.justice.gov/archives/opcl/page/file/844481/download>.
- ResearchDatagov, accessed October 6, 2022, <https://www.researchdatagov.org/>.
- Secure Research Data Network Act, S.4200, 117th Cong., 2022, <https://www.congress.gov/>

[bill/117th-congress/senate-bill/4200/text?r=23&s=1](https://www.congress.gov/bills/117/senate-bills/4200/text?r=23&s=1).

- Snyder, Belinda and Jeffrey W. Thomas, “GOGOs, GOCOs, and FFRDCs... Oh My,” *Federal Laboratory Consortium for Technology Transfer*, 2019, <https://webcasts.zilretta.com/flc/courses/43708/documents/64956>.
- South Carolina Department of Commerce, “South Carolina Coordinating Council for Workforce Development,” accessed October 6, 2022, <https://www.sccommerce.com/sc-coordinating-council-workforce-development>.
- South Carolina Revenue and Fiscal Affairs Office, “Health Care: Information on Utilizations and Professions,” accessed October 6, 2022, <https://rfa.sc.gov/data-research/healthcare>.
- Stokes, Peter, “The ‘Five Safes’–Data Privacy at ONS,” *U.K. Office for National Statistics* (blog), January 27, 2017, <https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/>.
- Title 13—Census, <https://uscode.house.gov/view.xhtml?path=/prelim@title13&edition=prelim>.
- Title 26—Internal Revenue Code, <https://uscode.house.gov/view.xhtml?path=/prelim@title26&edition=prelim>.
- United Nations Global Working Group on Big Data, “Privacy Preserving Techniques Task Team of the UN Committee of Experts on Big Data and Data Science for Official Statistics,” accessed October 6, 2022, <https://unstats.un.org/bigdata/task-teams/privacy/index.cshtml>.
- United Nations Global Working Group on Big Data, Privacy Preserving Techniques Task Team, “UN Handbook on Privacy-Preserving Computation Techniques,” accessed October 6, 2022, <https://unstats.un.org/bigdata/task-teams/privacy/UN%20Handbook%20for%20Privacy-Preserving%20Techniques.pdf>.
- United States Copyright Office, “Section 1201 Rulemaking: Eighth Triennial Proceeding to Determine Exemptions to the Prohibition on Circumvention,” October 2021, [https://cdn.loc.gov/copyright/1201/2021/2021\\_Section\\_1201\\_Registers\\_Recommendation.pdf](https://cdn.loc.gov/copyright/1201/2021/2021_Section_1201_Registers_Recommendation.pdf).
- U.S. Chamber of Commerce Foundation, “Jobs and Employment Data Exchange (JEDx),” accessed October 6, 2022, <https://www.uschamberfoundation.org/JEDx>.
- Workforce Innovation and Opportunity Act, Pub. L. No. 113–128, 2014, <https://www.congress.gov/bill/113th-congress/house-bill/803/text>.

# Acronyms

Acronym	Definition
ACDEB	Advisory Committee on Data for Evidence Building
ADC	America's DataHub Consortium
AERDF	Advanced Education Research and Development Fund
API	Application Programming Interface
ATI	Advanced Technology International
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
CARES	Coronavirus Aid, Relief, and. Economic Security
CCWD	Coordinating Council for Workforce Development
CDO	Chief Data Officer
CIPSEA	Confidential Information Protection and Statistical Efficiency Act
DOI	Digital Object Identifier
DOL	Department of Labor
DP	differential privacy
EO	Evaluation Officer
EPA	Environmental Protection Agency
ERS	Economic Research Service
FACA	Federal Advisory Committee Act
FAIR	Find, Access, Interoperate, and Reuse
FCSM	Federal Committee on Statistical Methodology
FedRAMP	Federal Risk and Authorization Management Program
FFRDC	Federally Funded Research and Development Center
FRN	Federal Register Notice
FSRDC	Federal Statistical Research Data Center
FTI	Federal Tax Information
FY	Fiscal Year
ICPSR	Inter-university Consortium for Political and Social Research
ICSP	Interagency Council on Statistical Policy
IIJA	Infrastructure Investment and Jobs Act
IRS	Internal Revenue Service

Acronym	Definition
JEDx	Jobs and Employment Data Exchange
KPI	key performance indicator
LEHD	Longitudinal Employer-Household Dynamics
MOU	memorandum of understanding
MWC	Midwest Collaborative
NASWA	National Association of State Workforce Agencies
NCATS	National Center for Advancing Translational Sciences
NCHS	National Center for Health Statistics
NCSES	National Center for Science and Engineering Statistics
NIH	National Institutes of Health
NORC	National Opinion Research Center
NSDS	National Secure Data Service
NSF	National Science Foundation
NVSS	National Vital Statistics System
OMB	Office of Management and Budget
OPEN	Open, Public, Electronic and Necessary
OUSEA	Office of the Under Secretary for Economic Affairs
PFAS	per- and polyfluoroalkyl substances
PMO	Project Management Office
PPRL	privacy-preserving record linkage
PPT	privacy-preserving technology
R&D	research and development
SAP	Standard Application Process
SMC	Secure Multiparty Computation
SO	Statistical Official
SOI	Statistics of Income
SWIS	State Wage Interchange System
TMF	Technology Modernization Fund
UI	unemployment insurance
USDA	United States Department of Agriculture
WIAC	Workforce Information Advisory Council
WIOA	Workforce Innovation and Opportunity Act



**Advisory Committee on Data  
for Evidence Building:  
Year 2 Report**

[Visit the ACDEB website](#)

