

Report 5: A Blueprint for Recommended Services to be Offered by the National Secure Data Service for Federal Policymakers

ADC EPS 24-N5
February 27, 2026

Prepared by: Esther C. Nolton, Anand Desai, Joshua Hawley, Cynthia Phillips, Mary Van Allen
Presented to: May Aydin; National Center for Science and Engineering Statistics



Contents

Introduction	1
Role of the Federal Data Ecosystem for Evidence-Based Policymaking.....	1
Foundation of the National Secure Data Service	2
A Human-Centered Design Approach to Understand Policy User Needs	2
Recommended Services for the NSDS	4
Discovery and Navigation.....	4
Discovery and Navigation Recommendations.....	4
Potential Discovery and Navigation Services.....	4
Existing Discovery and Navigation Examples	5
Shared Services and Resources	5
Shared Services and Resources Recommendations	5
Potential Shared Services and Resources Services	6
Existing Shared Services and Resources Examples.....	6
Access and Linkage Infrastructure	6
Access and Linkage Infrastructure Recommendations	7
Potential Access and Linkage Infrastructure Services	7
Existing Access and Linkage Infrastructure Examples.....	7
Rate-Limiting Factors for Adoption and Use of the NSDS	7
Relevant Existing Examples	9
Summary of Recommended Services for NSDS 1.0	9
Adoption and Use Begins with Strategic Communications	12
Limitations and Opportunities for Future Research	12
Conclusions	13

America’s DataHub Consortium (ADC), a public-private partnership, implements research opportunities that support the strategic objectives of the National Center for Science and Engineering Statistics (NCSES) within the U.S. National Science Foundation (NSF). These results document research funded through ADC and are being shared to inform interested parties of ongoing activities and to encourage further discussion. Any opinions, findings, conclusions, or recommendations expressed above do not necessarily reflect the views of NCSES or NSF. Please send questions to ncsesweb@nsf.gov. This product has been reviewed for unauthorized disclosure of confidential information under DRN NCSES-DRN25-071.

Introduction

This report serves as a synthesis of preceding Reports 1-4¹ and a record of our findings regarding the services that would be useful for generating evidence for federal policy stakeholders. Rather than develop a new blueprint for these services, we have used the vision of the National Secure Data Service (NSDS)² and embedded our recommended services in the three layers: Discovery and Navigation, Shared Services and Resources, and Access and Linkage Infrastructure. Additionally, we discuss the factors that may limit the rate at which the NSDS will be adopted and used for policy purposes, given the unique contexts and needs policymakers have in order to interact with and use evidence to inform policy decisions. We close with an emphasis on the role of strategic communications to facilitate adoption and use of the NSDS, and other future opportunities to enhance sustainable implementation.

Role of the Federal Data Ecosystem for Evidence-Based Policymaking

Evidence-based policymaking that incorporates observation, testimony, precedents, knowledge, data, narratives, argumentation, analysis, and synthesis is a multifaceted idea with a long history recorded in many ancient texts³. The coinage of the modern-day term *evidence-based* is from an editorial by Guyatt (1991)⁴ where evidence-based medicine (EBM) was first presented as bringing a new approach to medical practice. EBM rejected intuition and clinical experience as credible bases for clinical decision making. This rejection was possible because there existed systematic and rigorous evidence drawn from randomized controlled trials (RCT) for assuring the efficacy of treatments and therapies. However, randomized controlled trials in the public policy context are likely to be infeasible, illegal, or immoral and therefore not an option for generating data as a basis for evidence for informing public policymaking.

What determines evidence for policymaking does not have a simple answer. **Evidence for policymaking is context dependent.** An evidence-based policy claim is the product of an argument in which the role played by evidence is contested. Hence, the data which are the basis of the knowledge that constitutes the evidence must fit the policy purpose. To achieve that, the policy issue must first be formulated as a policy problem that lends itself to resolution through evidence, which is derived from data via a logical argument. The data alone, while necessary, are not sufficient for generating the evidence.

An alternative approach to seeking evidence to inform policymaking is to follow a systematic process using information and arguments to generate the evidentiary basis (knowledge and argument) for a policy position or recommendation. It is in this context that we discuss the role of the federal data ecosystem overlaid by services to generate evidence for use by federal policymakers in their policymaking. Congress has provided support for evidence-based policymaking through the passage of the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act)⁵ and the 2022 CHIPS and Science Act⁶. **For evidence-based policymaking to become an integral part of the culture of the public sector, it is important for the U.S. Office of Management and Budget (OMB) together with agency leaders to continue to provide leadership, institutional support, and the resources for bolstering the federal data ecosystem and related services, including the NSDS.**

¹ <https://www.americasdatahub.org/adc-final-reports-eps-24-n5/>

² <https://nces.nsf.gov/initiatives/national-secure-data-service-demo/vision-for-a-future-nsds>

³ For example, The *Analects* of Confucius (China), Talmudic Writings (Middle East), Aristotle's *Politics* (Greece), Kautilya's *Arthashastra* (India) The Code of Hammurabi (Babylon)

⁴ Guyatt GH. Editorial: Evidence-based medicine. *ACP J Club*. 1991;114(2):16. doi: 10.7326/ACPJC-1991-114-2-A16

⁵ Foundations for Evidence-Based Policymaking Act of 2018, Pub. L. No. 115-435, 132 Stat. 5529

⁶ <https://www.congress.gov/bill/117th-congress/house-bill/4346>

This data ecosystem has always been used by a broad range of users such as researchers, journalists, and policy analysts to meet their professional information needs as well as by the general public both in the US and abroad. Although widely used, the **federal data ecosystem is a complex system** that is not readily accessible and is difficult to navigate. An OMB memorandum from 2019 (M-19-23)⁷ states that evidence is needed in order to “...make critical decisions about program operations, policy, and regulations, and to gain visibility into the impact of resource allocation on achieving program objectives.” Putting into practice the government-wide use of evidence-based decision making will require OMB to create operational policies and guidance on communicating and integrating the NSDS efficiently and effectively.

Foundation of the National Secure Data Service

To implement the Evidence Act, OMB memoranda provided guidance and supported the creation of multiple interagency committees. The funding for the NSDS Demonstration projects came through the 2022 CHIPS and Science Act. The NSDS is “envisioned as a set of shared services and a government-wide data linkage and access infrastructure to support evidence building.”⁸ Several studies, including this one, are different from the technology-focused efforts in that they respond to recent recommendations by the Advisory Committee on Data for Evidence Building (ACDEB)⁹ for the NSDS to elevate the importance of capacity building. The NSDS provides a unique opportunity to design services for specific categories of users, enhance the user experience, and meet their data needs efficiently and effectively.

The NSDS does more than provide access to data in the federal data ecosystem. It has various functions such as data concierge services, a capacity building center, communities of practice, toolkits, and the data usage platform, all of which can offer services for evidence building from the data. This project is an initial effort to provide insights into the types of services that federal policy stakeholders could benefit from in the early development of the NSDS and **creates a blueprint for these services that can begin to facilitate access and enhance the use of the federal data ecosystem** to support the evidentiary needs of federal policymakers. This summative blueprint for services is based on findings shared in Reports 1, 2, 3, and 4 from our literature review, policy case studies, and engagement with a panel of policy stakeholders, who included evaluators, policy analysts, policy influencers, and policymakers.

A Human-Centered Design Approach to Understand Policy User Needs

The federal data ecosystem together with the NSDS is a complex sociotechnical system in which the services for federal policymakers are designed to support the undertaking of two independent evidence journeys: one journey from data to evidence and the other from a policy issue to evidence. The **data journey** is from *data*¹⁰ in the federal data ecosystem to *evidence* that can inform policymaking. The **policy journey** follows a *policy process* from identifying a *policy issue* to identifying and accessing the policy-relevant data to generate the necessary

⁷ <https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf>

⁸ Madray, H. (2024). Towards a National Secure Data Service: Two years of Progress from the NSDS Demonstration Project. National Center for Science and Engineering Statistics. https://players.brightcove.net/679256133001/NkgrDczuol_default/index.html?videoId=636590875011

⁹ <https://www.bea.gov/sites/default/files/2022-10/acdeb-year-2-report.pdf>

¹⁰ In this report, when we use the term “data”, we mean information in that the data ecosystem has an infrastructure for organizing the data and transparency and trust in an institutional environment that inform the data user about fitness for use of the data including sources, methods for collecting, storing, retrieving, processing, and analyzing the data.

evidence. Neither journey follows a linear series of steps leading to evidence necessary for policymaking.

The data journey to evidence, however, is a simpler path that begins with data that have form and structure that yield the information that exists in the federal data ecosystem. This information, when contextualized, is knowledge that makes sense of the information. The next step is for the knowledge to be accompanied by value-based arguments with warrants, rebuttals, and backing for that knowledge to be used as evidence to make policy claims and counterclaims.

The policy journey to evidence is more convoluted. First, there are multiple, not necessarily independent, types of evidence required at different stages of the policy process. Evidence is necessary to determine whether there is a need for a new policy to address a problem or would existing policies be sufficient to address them. If it is determined that a new policy is necessary, then there is a second need for evidence to assess how best to frame the evidence to seek the ensuing policy solution. Evidence is necessary to determine the framing, for instance, what form should the potential approach take? Should it be legal, political, administrative or something entirely different? Once that framing is established, there is a third set of questions regarding the policy remedies and the evidence necessary to explore potentially competing solutions. Finally, the policymakers need evidence to select among the policy options to find those that seem promising. “Promising” is a value-laden term for which criteria must be established, and indicators have to be developed to assess the possible options. It quickly becomes clear that the evidence required by the policymaker might not always be found in the federal data system, and even if the information is to be found there, the (data governance or administrative) services required to arrive at that information may not be within the scope of NSDS.

Our data sources—literature review, case studies, I3P members—yielded conceptual, practical, procedural, and experiential insights regarding the services NSDS could provide and those that would be necessary to support NSDS to ensure its success. We raise these issues here because although NSDS can provide a broad range of services that can help the data users with access, navigation, and discovery, there are also other services related to infrastructure that must be in place for NSDS to be successful.

Our findings suggest that the NSDS will, initially, be best equipped to provide access to and use of information in the federal data ecosystem. In addition to providing access, the services that facilitate data use for evidence-based policymaking will likely require institutional resources and senior leadership support for formulating policies and making choices among policy options.

Our project team conducted a human-centered design (HCD) study with the goal of understanding unique user journeys, needs, and expectations of the NSDS for informing policy purposes. The lessons we draw here are from three sources of knowledge: a literature review, engagement with an expert panel, and case studies. Our literature review touched upon multiple literatures including those on policymaking, decision making, knowledge management, evidence, argumentation, and complexity to gain insight into evidence-based decision making and how policymakers seek evidence for making policy. This review provided a conceptual foundation for our work.

We constituted a panel of policy analysts, policymakers, and evaluators who were familiar with the Evidence Act and its requirements for evidence-based policymaking. We used human centered design (HCD), which is a participatory, iterative, systems thinking approach for our engagement with the panel of policy stakeholders. At some time during their career, each

member of the panel worked with federal data. The panel members were well-aware of the challenges of obtaining and using data for policy purposes and had also developed solutions to address those challenges.

Our individual and collective engagement with members of the panel gave us insights into the types of services that could be designed and offered by the NSDS specifically to support evidence-based policymaking. This engagement also gave us a practical context in which to locate some of our findings from our review of the literature; the literature informs and influences public policymaking. The synthesis of the knowledge from this engagement offered an experiential foundation for the types of services needed to facilitate data access and data use for evidence-based policymaking.

Findings from our engagement with the panelists were bolstered by case studies that we developed by interviewing and working with policy analysts on their journey from an unstructured policy problem to delivering evidence to policymakers for the development of policy options and recommendations. We gained procedural and process-oriented knowledge from developing the case studies. The policy analysts served as intermediaries who translated the policy needs of the policymakers into information needs that could be met by the information in the federal data ecosystem. In addition to navigation and discovery, many of the services such as toolkits, the Capacity Building Center, the sandbox, and secure linkages, currently envisioned in the NSDS model¹¹ will meet policy stakeholder needs.

Recommended Services for the NSDS

As mentioned above, rather than develop a new blueprint for the services to support the needs of federal policy makers, we aligned to the three proposed layers of the NSDS, namely: Discovery and Navigation, Shared Services and Resources and Access and Linkage Infrastructure as the existing framework within which to embed the proposed services.

Discovery and Navigation

Our interviews with members of our expert panel revealed that access to data from the federal data ecosystem can be cumbersome but is feasible for experienced users. Here the proposed data concierge has an important role to play for the data to be readily accessible and useful to all users, regardless of skill level. Discovery and navigation are both important services for any user of the federal data ecosystem. Discovery is related to purpose and use. The ability to ask productive questions is an important service that the NSDS can provide to help initiate the evidence building process. Questions are often more difficult to articulate, especially in the policy context where problems are typically poorly formulated and might need help from subject matter experts for proper definition of the need. Hence, **perhaps the most important service that the NSDS can provide is to help the policy stakeholder define the data need(s)** with respect to the federal data ecosystem. In comparison, navigating a well-organized data system is relatively simple, at least when the data needs are known. **Good metadata, which are data about the data, improve the efficiency of the search and are essential for access to data.**

Discovery and Navigation Recommendations

DN1. The data concierge should serve as a knowledge broker and partner that facilitates data discovery and navigation of the data ecosystem. Provision of metadata, including comprehensive data catalogs and dictionaries that provide an inventory and definitions

¹¹ <https://nces.nsf.gov/initiatives/national-secure-data-service-demo/vision-for-a-future-nsds>

of available data as well as a data validation process are essential to assure transparency and confidence in the data for policymaking purposes.

Potential Discovery and Navigation Services

- An initial service that identifies the user as a policy stakeholder and routes the individual to a dedicated concierge designed to address queries regarding policymaking.
- A screening interview (e.g., with a chatbot or human) with the user to identify the policy area to guide the user to additional resources aligned with their data query.
- Identify and maintain metadata dictionaries that could point to other associated and complementary data that could be used in conjunction with the requested data.
- Provide a tiered service with multiple pathways and routing intelligence based on user sophistication levels, with different concierge levels and service models appropriate for varying technical comfort levels and time constraints.

Existing Discovery and Navigation Examples

- [National Archive of Criminal Justice Data](#) (University of Michigan)
- [Data Governance Transformation Project](#) (Massive Data Institute)
- [Community Data Platform](#) (mySidewalk)
- [World Bank Open Data](#) (World Bank Group)
- Proprietary and open access longitudinal data across the research, development, and innovation enterprise (e.g., [Clarivate](#), [OpenAlex](#))

Shared Services and Resources

Policies are complex and their evidentiary needs can rarely be met by data from a single database or agency. This need for related or adjacent data suggests a set of services that can guide the user to other data and data sources that might be used in conjunction with data explicitly sought by the user. Providing access to data is perhaps a relatively easy service to offer. **This service is not unique to any specific type of user; however, for policymaking, the ability to efficiently identify, blend, and integrate data from multiple sources becomes an important service capability.** Moreover, the ability to provide access to the data, framed by the appropriate data governance and security considerations is paramount. Also, providing the support and expertise of subject and data experts to identify data sources most relevant for the policy user will encourage confidence in the NSDS support. With the rapid evolution of generative artificial intelligence (AI) chatbots, it is not only possible to automate responses to well-defined queries, but to also rapidly respond to open-ended questions about potential sources and data variables to meet the users' needs.

Members in our expert panel relied upon a broad network of subject matter and data experts to help them in their discovery and navigation of the landscape of policy-relevant data. Working on the case studies further emphasized the need for and importance of such communities of practice for policy analysts seeking information quickly within a short policy window. **Such groups are built over a career, but the NSDS can expedite the process by institutionalizing issue-based networks.** Our panel members emphasized the need for data literacy among policymakers. Their observations are echoed in the policy literature where timeliness and presenting audience relevant information in the right format are important for evidence-based policymaking. Such skills can be taught, and training could be included in the set of services provided in the proposed Capacity Building Center of the NSDS.

Shared Services and Resources Recommendations

- SR1. Incorporate advanced tools for locating and disaggregating relevant data. Identify accessible staff and partners, including from nonprofit and academic organizations with knowledge of the federal data ecosystem, to support exploration, discovery, and access.
- SR2. Facilitate the identification of practitioners and experts to build communities of practice and knowledge that can be accessed by policy analysts for guidance on finding relevant data, scholarly and gray literature, and policy precedents. Some of these services can be automated such as AI-enabled identification of extant reviews of the relevant literature.
- SR3. Provide training and capacity building in presenting the outputs of statistical and other analytical tools in formats accessible by and appropriate for the intended audience.
- SR4. Provide services that identify data sets and data used by other users asking similar policy questions.

Potential Shared Services and Resources Services

- Capacity building center
 - Maintain a repository of use cases where such data were used in the past.
 - Provide training for users to design queries that describe their data needs to make data retrieval and use more efficient.
 - Offer services to help users bring in data through secure linkages to examine the potential for integration and analysis in the sandbox.
 - Offer training to use novel analysis tools and techniques that are made available in the sandbox.
- Communities of practice
 - Provide lists of potential subject matters experts who can address policy relevant questions regarding the data request.
 - Support for creating communities of practice for users to accelerate their discovery without relying on extensive personal networks.
- Data usage platform
 - Identify frequently requested federal datasets to deliver immediate value to the time constrained user base while building institutional credibility.
 - Identify data from multiple sources that are commonly used together to facilitate joint use and data sharing.

Existing Shared Services and Resources Examples

- [Microsoft](#) or [Marriott](#) and other corporate organizations
- [Secure Query Service](#) (Massive Data Institute)
- [KYStats](#) (Kentucky Center for Statistics)
- [Analyst's Guide](#) and [Resources](#) to Federal Spending Data (USASpending.gov)
- State and local longitudinal data systems (e.g., Allegheny County [Children, Youth, and Families](#), DC [Education Through Employment Pathways](#), [Ohio Education Research Center](#), Virginia [Longitudinal Data System](#))

Access and Linkage Infrastructure

Policies are made at various levels of government. At any given level, the policy-relevant data are distributed across multiple agencies and include localized data from varying sources. Although our focus is on the federal data ecosystem, which is vast, the national data ecosystem, which includes state, local, tribal, and commercial data, is in fact an order of magnitude larger. **A federated architecture with distributed systems and local services will make it possible to provide more effective services in addition to enhancing the efficiency of data**

management. A potential additional advantage of providing a secure compute environment as a service is that it can accelerate data access and integration through secure linkages across agencies. Another advantage of having a secure compute environment is that it facilitates the creation of a sandbox, close to the source of the data where policy analysts can explore through the simulation of multiple models and their behaviors to understand the sensitivity to various assumptions and different types of data at various levels of granularity, aggregation and disaggregation.

Access and Linkage Infrastructure Recommendations

- AL1. Develop a services infrastructure that serves cross-agency evidence building needs. Embedding concierge services within federal agencies could be a first step in building such infrastructure.
- AL2. Offer a secure sandbox with dedicated live support for rapid prototyping and testing with data from multiple sources, including synthetic data. In addition to data, the sandbox can facilitate exploration by providing sophisticated analytical and visualization tools to examine the consequences of using different data models, data sources, and units of analysis.

Potential Access and Linkage Infrastructure Services

- Sandbox
 - Offer shared spaces for capacity building among communities of practice coming together to develop solutions.
 - Provide service to facilitate users bringing other data through secure linkages to examine potential policy solutions.
 - Offer alternative tools such as simulation models and training for exploring scenarios of potential policy outcome as alternatives to more traditional predictive analysis tools.
- Secure linkages
 - Offer secure standardized linkages to data that are commonly used for policy purposes and reside outside the federal data ecosystem.

Existing Access and Linkage Infrastructure Examples

- [Reference Architecture](#) and [Disclosure Avoidance Redaction Tool](#) (Community of Innovation)
- [Data Sources and Linking](#) (USASpending.gov)
- [Data Concierge Service](#) (Harvard FAS Informatics Group)
- [Equitable Value Explorer](#) (Gates Foundation)
- [Safeguarding the Entire Community in the U.S. Research Ecosystem](#) (SECURE Center)
- Upcoming [Indicators Explorer](#) (NCSES)

The Research and Development (R&D) and Analytics sandbox can provide policy stakeholders with the opportunity to simulate and explore the consequences of using different data models and different types of data to develop scenarios under different policy assumptions. Potentially, secure linkages to data within and outside the federal data ecosystem can also be established to supplement access to federal data. Both these capabilities, the sandbox and the linkages, can provide services in the secure computing environment of the Access and Linkage Infrastructure. Many of these services can generate evidence motivated by the kinds of decisions that have to be made and to identify or develop policy options. Access to such information broadens the notion of evidence in evidence-based policymaking. In the following

section we discuss the services related to both access to and use of the federal data ecosystem and the role of these services in evidence building.

Rate-Limiting Factors for Adoption and Use of the NSDS

Our panel also discussed a number of cultural and institutional factors that could enable and enhance the adoption and use of evidence for policymaking, which are beyond the scope of the NSDS but are important to the success of the NSDS. Many of the services necessary to mitigate these **rate-limiting factors require large-scale sociotechnical systems change** in the federal data ecosystem and are outside the scope of the NSDS. **However, lack of attention to these factors runs the risk of disabling the NSDS from becoming a useful tool in how policy decisions are informed by relevant and useful data.** For instance, the presumption of accessibility is yet to be fully realized. **In the absence of resources and a concerted effort by OMB and data owners, it is likely to remain an issue plaguing seamless access to and use of the federal data ecosystem.**

Many of these factors are rate limiting because of a lack of long-awaited OMB guidance on data governance and regulations on data access (i.e., despite being stated objectives of the Interagency Council on Statistical Policy¹² to both launch the NSDS 1.0 with clarifying linkages and provide expertise to data regulations' development), among other issues, some of which require establishing interagency councils to facilitate and promote collaboration among data providers in different agencies and units of the government.

In discussing the solutions to the challenges that our panel members encountered, they also mentioned factors that limited the efficacy of their solutions. **Many of the limitations they encountered in accessing and using data from the federal data ecosystem were administrative, procedural, or legal in nature.** They could not be mitigated by any service that the NSDS could offer because of their structural nature. For example, the barriers encountered pertained to the localized nature of governance structures which limited their ability to quickly access and blend or integrate data from multiple sources with different owners. Their lack of ability to obtain these data for addressing time-sensitive policy issues was tantamount to them not being able to conduct their analyses and address problems because of the lack of evidence. Another issue is the lack of standardization of common data access and use policies, which had similar adverse effects on conducting analyses for time sensitive policy changes. The ubiquity of evidence-based policymaking is a shift that can only be achieved through long-term changes in culture, habits, and practices in the contextual environment in which NSDS exists. **OMB is in a unique position to serve as a critical change agent in providing regulatory guidance on common data access and use policies.**

Lack of timely access to data limited policy stakeholders' ability to inform legislators or agency staff to help build advocacy coalitions to promote policy changes or systemwide cultural change. Lack of standing councils of interagency leaders also hindered the sharing of data to inform collaboration and coordination across agencies to achieve mutually beneficial objectives. Cross-agency infrastructure could facilitate data sharing and use but they are not services that can be provided within the NSDS. However, there are services that the NSDS could provide in this context. For example, **the sandbox in the secure compute environment can provide policy analysts with a space for cross-agency collaboration on rapid prototyping using real or synthetic data.** They could also collaborate in the sandbox on developing comprehensive data catalogs and validation processes that could be made available in the Capacity Building Center.

¹² <https://statspolicy.gov/about/#icsp>

These rate-limiting factors are recognized concerns in other contexts, and several efforts attempt to mitigate their effect on data generation and use in policy contexts. While none reaches the scale and complexity required for the federal data ecosystem, some existing examples could be useful to consider as part of the NSDS or as ancillary services that support the ultimate use and effectiveness of the NSDS. Large-scale sociotechnical systems change is already complex. **Greater awareness of these rate-limiting factors helps manage expectations regarding the speed at which we can expect change to occur for the NSDS.**

Relevant Existing Examples

- [Administrative Data Research UK](#) (UK Research and Innovation)
- [Learning Agenda Toolkit](#) (Grant Thornton/Office of Evaluation Science)
- [Creating a Learning Agenda for Systems Change](#) (Public Health Learning Network)
- [Data Skills for Congress](#) (USAFacts/UC Berkeley)
- [Data Literacy for Leaders](#) (Public Service Leadership Institute)
- [Data Literacy Workshop Series](#) (Academic Senate for California Community Colleges)
- [Federal Interagency Council on Economic Mobility](#) (HHS ASPE)
- Connect [Gold Standard Science](#) and other standards of scientific rigor to the business of evidence-building
- Technology adoption cohorts (e.g., [TECH Clean California](#), [SBA](#), [NIST](#))
- [JEDx](#) (U.S. Chamber of Commerce)
- Stories of data failures leading to public demand (e.g., [UPWARD](#) in California, [CMS Blue Button](#), IRS [Direct File](#), Census [Differential Privacy](#))

Summary of Recommended Services for NSDS 1.0

With the increasing sophistication of agentic AI, it is now possible to respond to natural language queries and provide navigation services without human intervention. Queries regarding discovery and other services can be partially addressed using AI before being handed over to human service providers.

When a policy stakeholder first encounters NSDS, speed in identifying the user as such is perhaps most important. Many experienced policy stakeholders will likely know the data and their sources within the federal data ecosystem and the data concierge should quickly facilitate access. All other policy stakeholders would be directed to an interview (AI or human) where their data and service needs will be assessed. If in the interview, it is determined that they only need data, it is a simple request that can be readily satisfied. Any other data and service need will need the data concierge to create a pathway to the data and the associated services.

Over time, patterns will appear in the type of data and services sought, but for NSDS 1.0 most of the combinations of services and data will be unique and the sequence in which the services and data are needed will vary. For instance, users might need a toolkit and then the data to experiment with in the sandbox before determining whether the data meet their specific purpose. They might seek out subject matter experts in an issue driven community of practice, determine that they need some training from the Capacity Building Center before requesting additional data.

Hence, although the path to satisfying their data and service needs might vary, we have identified a number of services that will be needed by policy makers. For ease of access, we summarize these services and present them in one location in the table below

Systems Layer	Prioritized Strategies	Potential Services
Proposed NSDS Layers		
Discovery and Navigation	<ul style="list-style-type: none"> ● Position data concierge as a knowledge broker and partner ● Embed concierge services within agencies ● Make human experts available for special support ● Build comprehensive data dictionaries that provide an inventory and definitions of available data 	<ul style="list-style-type: none"> ● An initial service that identifies the user as a policy stakeholder and routes the individual to a dedicated concierge designed to address queries regarding policymaking. ● A screening interview (e.g., with a chatbot or human) with the user to identify the policy area to guide the user to additional resources aligned with their data query. ● Identify and maintain metadata dictionaries that could point to other associated and complementary data that could be used in conjunction with the requested data. ● Provide a tiered service with multiple pathways and routing intelligence based on user sophistication levels, with different concierge levels and service models appropriate for varying technical comfort levels and time constraints.
Shared Services and Resources	<ul style="list-style-type: none"> ● Incorporate advanced tools for locating and disaggregating relevant data ● Identify accessible staff and partners, including from nonprofit and academic organizations, to support exploration, discovery, and access ● Build communities of practice to promote knowledge exchange 	<ul style="list-style-type: none"> ● Capacity building center <ul style="list-style-type: none"> ○ Maintain a repository of use cases where such data were used in the past. ○ Provide training for users to design queries that describe their data needs to make data retrieval and use more efficient. ○ Offer services to help users bring in data through secure linkages to examine the potential for integration and analysis in the sandbox. ○ Offer training to use novel analysis tools and techniques that are made available in the sandbox. ● Communities of practice <ul style="list-style-type: none"> ○ Provide lists of potential subject matters experts who can address policy relevant questions regarding the data request. ○ Support for creating communities of practice for users to accelerate their discovery without relying on extensive personal networks. ● Data usage platform <ul style="list-style-type: none"> ○ Identify frequently requested federal datasets to deliver immediate value to the time constrained user base while building institutional credibility. ○ Identify data from multiple sources that are commonly used together to facilitate joint use and data sharing.

<p>Access and Linkage Infrastructure</p>	<ul style="list-style-type: none"> ● Develop an infrastructure that serves cross-agency evidence building needs ● Ensure data are validated, focused, and relevant ● Improve readiness of data to be integrated and used by AI tools ● Streamline the process of obtaining data sharing agreements between agencies 	<ul style="list-style-type: none"> ● Sandbox <ul style="list-style-type: none"> ○ Offer shared spaces for capacity building among communities of practice coming together to develop solutions ○ Provide service to facilitate users bringing other data through secure linkages to examine potential policy solutions. ○ Offer alternative tools such as simulation models and training for exploring scenarios of potential policy outcome as alternatives to more traditional predictive analysis tools. ● Secure linkages <ul style="list-style-type: none"> ○ Offer secure standardized linkages to data that are commonly used for policy purposes and reside outside the federal data ecosystem.
<p>Rate-Limiting Factors</p>		
<p>Habits and Practices</p>	<ul style="list-style-type: none"> ● Offer safe sandbox with dedicated live support for rapid prototyping and testing with synthetic data ● Build robust data catalogs and validation processes 	<ul style="list-style-type: none"> ● Sandbox <ul style="list-style-type: none"> ○ Enhance sandbox capabilities to promote collaborative learning ● Communities of Practice <ul style="list-style-type: none"> ○ Provide institutional support by creating interagency councils and facilitate development of communities of practice
<p>Data Literacy</p>	<ul style="list-style-type: none"> ● Bridge the gap between policy and evidence ● Focus on congressional engagement as champions ● Enhance skills in formulating clear and feasible policy questions ● Establish issue-based councils for cross-agency collaboration and coordination 	<ul style="list-style-type: none"> ● The Capacity Building Center <ul style="list-style-type: none"> ○ Offer training to Communities of Practice that can work collaboratively in the Sandbox to explore common data interests and solutions. ○ Provide services to enhance data literacy through the creation of toolkits (e.g., guiding principles, checklists, policy rubrics). ● Communities of Practice <ul style="list-style-type: none"> ○ Institute interagency councils for promoting collaboration across data owners outside NSDS that can support the creation of issue-based communities of practice
<p>Culture and Leadership</p>	<ul style="list-style-type: none"> ● Communicate clearly the awareness and value to range of users ● Endorsement and ownership from top-down to promote systems-wide culture change 	<ul style="list-style-type: none"> ● Most of the institutional, governance, administrative, and legal services will be offered outside NSDS through data owners with OMB leadership.

Adoption and Use Begins with Strategic Communications

Building the NSDS is not enough. Throughout our engagements with the panel, the members expressed enthusiasm and support for the NSDS becoming a helpful tool for enhancing the utility and use of the federal data ecosystem. For this reality to become a success, they emphasized the need for a champion who would endorse and take ownership of the NSDS implementation—for which they nominated OMB as the best candidate for central coordination and communication. Together, **OMB and the senior leadership of federal agencies need to explicitly champion the NSDS** for the widespread use of the federal data ecosystem.

This communication framework draws on the Prosci® ADKAR model¹³ to facilitate adoption and use of NSDS. The components of the model are to create **Awareness** of the data ecosystem and the NSDS; clarify the value of the NSDS to create a **Desire** for its use; provide **Knowledge** to help potential users understand what it can do; facilitate adoption by offering training to enhance potential users’ **Ability** to use the NSDS; and finally, provide **Reinforcement** through clear, authoritative guidance from OMB about adoption and use of the NSDS.

ADKAR Step	Suggested Activities or Communication Strategies
Awareness	<ul style="list-style-type: none"> ● Develop a flagship, bipartisan pilot use case¹⁴ that demonstrates the functionality of the NSDS (e.g., including use cases at the state and local government levels). ● Publicize the process, results, and policy impact widely to establish the NSDS as a state-of-the-art solution to navigate and access data. ● Lead communications with emerging or early-adoption use cases, not architecture, based on identified priority questions from diverse policy users.
Desire	<ul style="list-style-type: none"> ● Clarify the value proposition by adopting proactive framing of the NSDS as a navigation and concierge service to help users find, access, and use data for policy purposes. ● Emphasize problems the NSDS will address and provide practical, question-specific demonstrations of what the NSDS will enable policy stakeholders to answer.
Knowledge	<ul style="list-style-type: none"> ● Build evidence of cross-agency infrastructure value by demonstrating how the NSDS serves interagency evidence-building needs. ● Show real examples of problems that can and cannot be solved with coordination and implementable solutions.
Ability	<ul style="list-style-type: none"> ● Equip and deploy distributed, local champions at each agency to facilitate adoption. ● Enable hands-on experiences through low-barrier sandboxes and experimental environments for training.
Reinforcement	<ul style="list-style-type: none"> ● Provide clear, authoritative operational guidance in collaboration with OMB.

Limitations and Opportunities for Future Research

This project was conducted by a small team in a year of significant federal change. The presented recommendations merely scratch the surface of the various ways in which the NSDS can enhance the access to and use of the federal data ecosystem through policy communities. Even without the NSDS services, the data ecosystem is a useful and used resource. Our project

¹³ <https://www.prosci.com/methodology/adkar>

¹⁴ A [pilot use case](#) allows a complex digital transformation project to be tested for feasibility and scalability. Observed pain points can be used to inform refinements while successes may be communicated as promise and value of the NSDS.

was focused on policy stakeholders who are a narrow but important slice of the full range of current and potential users of the NSDS. Our group of key informants was a robust selection of experienced leaders but could be considered relatively small when compared to the full range of services that the NSDS could provide policy stakeholders.

For evidence-based policymaking to become universal, it will need many more services and resources to generate the evidence needed to address the need. As the NSDS is implemented, similar studies will have to continue to foster systematic identification and development of additional organically formed and emergent services. Many of these services deal with reducing barriers to the interoperability of multi-agency data, the lack of seamless governance structures and standardization of data structures, the difficulties in achieving full transparency in an environment where the technology is evolving rapidly and the databases can grow exponentially.

Conclusions

The users of NSDS span a wide range from naïve data users to seasoned power users. To serve the full set of needs, the services will likely range from responses to open-ended questions about data options to learn about a topic of interest to well-defined queries for retrieving specific variables as stipulated by the user from one or more databases in the data ecosystem. Identification of the type of user upon entry into the NSDS should be a screening feature located at the entry point of the NSDS. A second-level filter for identifying individuals who require support to generate evidence for policymaking should also be an early screening activity. Once these screenings have been done, the services we have suggested are likely to be more effective.

We have made a distinction between the use of *evidence in* and *evidence for* evidence-based decision making. Using evidence in policymaking is necessary to determine whether there is a sound basis for making a policy claim. On the other hand, using evidence for policymaking is motivated by the policymaker in terms of the person's need to seek evidence, for instance, to choose among policy options. The NSDS can serve both those needs. Notably, the policy process requires rapidity and the NSDS can offer services that facilitate more efficient data access and use. The role of AI will be an important, potentially long-term, concern for the deployment of the NSDS. Using AI provides an appealing option, especially when resources are limited. In many instances, AI may be the preferred option. However, extensive testing of the output and services that AI provides will be essential to engender trust and attract users regardless of their attitude towards its use.

The NSDS 1.0 is yet to be released so the range of service needs that it will satisfy is not fully known. Regardless of the sophistication of the services provided, the existence of those services will yield more ideas about how they can be modified, improved upon, and supplemented or replaced. A successful NSDS could begin to attract a broader group of users who will bring additional unique needs for services. Once the system is operational, concerns regarding transparency, interoperability, governance, and administration will evolve and will need to be continually reviewed. New solutions and additional research will result in the need for hitherto unknown services regarding both access and use of the current federal data ecosystem or an enhanced, national data ecosystem. Some of these services may be beyond the scope of the NSDS but are essential for smooth access and ready use of the federal data ecosystem (e.g., services related to institutional arrangements, governance and legal structures, data privacy, and security). Despite the daunting complexity of the large-scale sociotechnical system that is being built, democratizing data by making them widely accessible is a worthwhile objective. The NSDS has the exciting potential to be a major step in that direction.