Report 2: Current State of Evidence Use by Federal Policy Stakeholders

ADC EPS 24-N5
September 25, 2025

Contents

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.

1. Introduction

This report presents findings from an examination of federal policymakers' data needs and challenges encountered in seeking evidence to inform policy decisions. These findings will inform the development of the future National Secure Data Service (NSDS). Through interviews with 12 senior federal officials, state government leaders, and policy experts, we identified critical gaps in the current data infrastructure and specific requirements for enhancing evidence-based policymaking across government.

1.1 Report Overview

1.1.1. Report Purpose

The purpose of this report is to provide an overview of our methods for primary information gathering and a summary of findings from our data collection efforts. We will further build on the concepts introduced in our first report, which provided the foundational background for the purpose and plan for user research. This report provides the preliminary findings from 12 key informant interviews (KIIs) plus an examination of one unique use case, focusing on the takeaways from all the interviews and the case study.

1.1.2. Organization of the Report

This report provides more details about the project impetus and framing of the current state of federal data use for policy purposes. We further clarify what we mean by federal data and the generation of a use case as an example of the current state of federal data use. We provide an overview of information gathering methods, including who participated in those activities and how we analyzed the data. We also provide a summary of findings, including takeaways from the expressed challenges across the 12 KIIs and the use case. These takeaways will help provide a better understanding for the mapping of challenges experienced by diverse federal data users. This mapping can signal the expectations of a future NSDS to meet user needs. We close this report with a preview of how these findings will be synthesized and integrated into user models and maps that will be presented in our next report.

1.2 Project Overview

The following sections provide an overview of the background, purpose, and underlying concepts that frame our approach.

1.2.1. Project Impetus

The NSDS was a concept originally introduced as part of the final report from the Commission on Evidence-Based Policymaking published in 2017.[1] Among many recommendations for the future NSDS, the Commission called for a comprehensive data service that would “model best practices for secure record linkage” across both federal administrative and statistical data. The Evidence Act of 2018, further established an Advisory Committee on Data for Evidence Building, which has led the effort in exploring the operationalization of the proposed recommendations for a NSDS.[2]

In 2022, the CHIPS and Science Act required that NSDS Demonstration (NSDS-D) projects be funded to support efforts to design and build a future NSDS that would be responsive and useful

[1] https://www2.census.gov/adrm/fesac/2017-12-15/Abraham-CEP-final-report.pdf
[2] https://datafoundation.org/news/data-sharing/39/39-A-Blueprint-for-Implementing-the-National-Secure-Data-Service

to data users.[3] There are currently 39 demonstration projects exploring different aspects of the demands and functions of a future NSDS. This project is completely focused on understanding the user needs of federal data from the perspective of these users to inform the anticipated uses and expectations of an effective NSDS.

1.2.2. Project Purpose and Goal

The original intent of this project was to answer two key focus questions:[4]

- What unique approaches and support do policy stakeholders need for data-driven decision making?
- How can a data concierge service support the needs of policy stakeholders?

Our team proposed a human-centered design approach to answering these questions and this is the first of a two-part report that will share findings that highlight the current state of federal data (and other related non-federal data sources) use. Primarily, in this report, we share what we have heard and interpreted as the key challenges and existing strategies when accessing and using federal data. In the next report, we will share more about proposed solutions and recommendations for the future state that can inform the future design and development of a robust NSDS.

1.2.3. Taking a Use-Inspired Approach to Research

This study takes a use-inspired approach to engage diverse stakeholders throughout the entire lifecycle of the research project.[5] A **Human-Centered Design (HCD)** approach involves seeking multiple expert perspectives to better understand the actual challenges and ideate about possible solutions to arrive at frameworks and prototypes that can be implemented and tested in practice.[6] This approach lends itself well to answering the two key focus questions for this project because it centers the user experience of various policy stakeholders and their needs of a future data concierge service and helps inform recommendations for a responsive and effective future NSDS.

Instead of beginning with a theoretical framework in mind, we began our exploration of the two key focus questions through grey and peer-reviewed literature to better understand the policy ecosystem. As reported in Report 1, we laid a foundation for the study by clarifying the policy process and evidence culture in the U.S. federal government, including the ecosystem of systems, people, and technology involved at all stages of the process. Next, it was necessary for us to hear directly from users about their experiences, which could not be captured in the literature review. Through the stories we heard from these informants, we are able to paint a fuller picture of the journey and experience of a user who might interact with the federal data ecosystem to answer policy questions, which may also involve non-federal data systems.

As a part of this study, we will also take a deep dive into one demonstrative use case that is focused on the experience of frontline analysts who interact with federal data for a specific policy purpose. This use case provides a practical example of how analysts might ultimately expect to use the future NSDS Data Concierge Service. It will make explicit “potential needs, gaps, successes, and challenges that are faced by federal policy stakeholders as they use federal data to support the development of policy recommendations of national significance”

---

[3] This includes data of all types including but limited to other systems and enclaves that house non-federal data (e.g., state/local/tribal government, private or commercial data systems).

[4] https://www.americasdatahub.org/wp-content/uploads/2024/06/ATT-1_Topic_Engaging-Policy-Stakeholders.pdf

[5] https://www.elgaronline.com/edcollchap-oa/book/9781803928289/book-part-9781803928289-21.xml

[6] https://designthinking.ideo.com/#design-thinking-context

(Report 1). It will, not in and of itself, produce recommendations for the future NSDS, but will offer insights into the current state of federal data use in a specific context and environment for developing a framework for a future NSDS.

1.2.4. Role of the Insight, Interpretation, and Innovation Panel

The **Insight, Interpretation, and Innovation Panel (I3P)** members were purposively selected[7] to represent diverse perspectives across the federal policy ecosystem. The I3P includes 12 senior-level professionals with 15-25 or more years of experience operating at the intersection of policy and data as described in Report 1 across four different stakeholder types (namely, policymakers, policy advisors, policy evaluators, and policy influencers). These individuals serve as "evidence brokers" who routinely translate between skills in federal data use for and policy decision making needs. They have each served in various leadership roles in multiple organizations that afforded them unique opportunities to interact with federal policies and data.

These members include seasoned professionals working at different government (state and federal) agencies (n=8) or non-governmental (e.g., academic or nonprofit) organizations (n=4). Each member holds one or more roles as a technical systems leader, senior agency leader, congressional or political liaison, or a state government official. The I3P represents a diverse group of senior policy experts and influencers who provide a wide coverage of perspectives about inquiry that serve different parts of the policy lifecycle and for different purposes. A full breakdown of I3P panel member demographics can be found in Appendix A. I3P members have provided important insights at every stage of this project. An objective of this report is to share findings from individual KIIs with each of the I3P members.

2. Methods

2.1. Approach to Key Informant Interviews

Between May and July 2025, all 12 I3P members participated in 60-120-minute KIIs with 2-3 interview facilitators. Members of the I3P were contacted via email to request and schedule individual interviews. In order to accommodate busy executive schedules, each KII was conducted over Microsoft Teams and was recorded on the same platform with each of the participants' consent. Each interview included 2-3 interviewers to support facilitation, notetaking, and member checking. The interviewers debriefed together after each interview. Transcriptions were retained for notetaking and analysis purposes, which were shared with the informants for review and recordkeeping. Participants were allowed to amend or delete any portion of their transcript. Participants were provided a series of questions that were part of a prepared interview guide (see full interview guide in Appendix B). Near the end of each interview, participants were provided an image of the proposed NSDS model[8] for reactions and interpretation of its interaction with or facilitation of their work. Informants also provided thoughts on the factors that would contribute to (un)successful NSDS implementation.

2.2. Analytic Methods

All interviews were conducted and transcribed using Microsoft Teams (Microsoft, 2025). Interview transcripts were processed initially by removing filler words and were provided to informants for recordkeeping and confirmation. Transcripts were then analyzed using a structured framework focusing on current data practices, challenges, proposed solutions, and user requirements. Claude.ai (Anthropic, 2025) was leveraged as a generative AI tool that offered rapid summaries and analyses. Cross-case pattern analysis identified common themes and divergent perspectives across different types of users and organizational contexts.

[7] https://pmc.ncbi.nlm.nih.gov/articles/PMC4012002/

[8] https://ncses.nsf.gov/initiatives/national-secure-data-service-demo/vision-for-a-future-nsds

Additionally, to verify the fidelity of summary themes generated by Claude.ai, it was specifically asked to provide evidence of the accuracy and veracity of its claims. The prompts used to generate and verify these themes are provided in Appendix C. An additional check of veracity of the AI summaries was conducted through careful review and discussion of the interviewers' own notes and perspectives with support from the raw interview transcripts.

All themes identified across all 12 KIIs were extracted and tabulated. Analysis involved systematic deductive coding of challenges, user needs, proposed solutions, and implementation considerations according to sociotechnical systems alignment (e.g., Social-Organizations, Social-Individuals, Technical-Tasks, and Technical-Technology) and mapping to the proposed NSDS model (e.g., Discovery and Navigation, Shared Services and Resources, Access and Linkage Infrastructure). For factors that were external to the NSDS (i.e., outside the federal data ecosystem but influences the data ecosystem), it was coded as an "External Cultural or Environmental Factor". Additionally, a Strengths, Weakness, Opportunities, Threats (SWOT) analysis was performed by also coding each theme as a Strength, Weakness, Opportunity, or Threat. The full codebook, including descriptions for each code, can be found in Appendix D.

To test and refine coding procedures to enhance trustworthiness, the codebook was co-created between two experts on the team who are very familiar with all three code constructs. One team member coded all of the themes and conducted interim peer debriefing[9] sessions with another team member and practiced reflexivity[10] exercises to check and recheck codes. Peer debriefing sessions were conducted by having a team member review 25 percent of the codes, which were randomly selected from batches of completed codes, and then holding meetings to discuss agreements, disagreements, and challenges. This process was repeated iteratively throughout the coding process to ensure continuous checking and alignment. Themes were organized in a table crosswalking NSDS model alignment with systems alignment. Themes were individually displayed in a font color matching that aligned with the SWOT analysis categories in order to present all three coding categories into one table. Themes that were repeated across multiple informants were consolidated into a single new theme.

During the KIIs, each informant shared several different user stories about their experiences interacting with data to answer policy questions. Stories varied in complexity, length, and comprehensiveness. However, a single illustrative vignette was teased out from each interview to highlight a specific example of the challenges experienced and current strategies employed by informants. Informants were provided an opportunity to review the illustrative vignette to ensure the language appropriately captured a highlight of their experience. All data from interviews were anonymized and are reported without any identifying information included, but we did use general descriptors to characterize the type of role, type of organization, and part of the policy lifecycle that is related to each vignette.

3. Findings

3.1. Type of Policy Stakeholders

A primary finding that emerged from the interviews was a description of the different types of data users and their role in the policy process. The four stakeholder types (e.g., Policymakers, Policy Advisors, Policy Influencers, and Policy Evaluators) that we originally described in Report 1 (and used to delineate policy perspectives when recruiting I3P members) were not actually mutually exclusive. These stakeholder types also did not include external staff (e.g., contractors) or program implementation staff, who both play a major role in the policy process.

[9] https://delvetool.com/blog/peerdebriefing
[10] https://delvetool.com/blog/reflexivity

Three persona types of distinct actors who use data and their needs (shown in Table 1) emerged from the interviews. Each informant shared experiences either serving as or collaborating with individuals across these three persona types. Together, their perspectives helped clarify the types of different actors involved in the federal data ecosystem. The descriptions of the three persona types shown in Table 1 became better descriptions of the role delineation of different actors who may interact with a future NSDS. Table 1 below shows a general summary of the different types of users and their needs of federal data:

Table 1. Types of Personas and Description of Policy Role

Persona Type	Policy Role
Congressional and Political Staff	Individuals who influence policy, but have no direct decision-making role
Agency Senior Policy Leadership	Senior advisors and evaluators who inform the agenda and make recommendations
Direct Data Users (e.g., Program Implementation Staff, Operational Staff, Analysts)	Agency staff or contractors who are: ● Subject matter experts and data providers ● Designers and technicians of data systems and infrastructure to support generation and use ● Data specialists who generate, analyze, and integrate evidence

In this discovery, as shown below in Figure 1, we realized that the individuals who are most likely to interact directly with a future NSDS are the Direct Data Users (e.g., Program Implementation Staff, Operational Staff, or Analysts), and not the two other persona types who have more prominent roles in formulating and influencing data strategy and guiding inquiry rather than working with the data. These Direct Data Users can be internal or external to the federal government but are most similar in their needs of and direct interactions with the data ecosystem to answer policy questions. This information regarding the personas and their level of interaction with the data ecosystem helps manage expectations about the various needs and priorities of potential NSDS users.

Figure 1. Proximity of Persona Type to the NSDS

3.2. Current State: Challenges and Pain Points

Although the informants provided many insights into proposed solutions (e.g., opportunities for interagency collaboration, suggestions for the design of a future NSDS), those are solutions oriented and mostly focus on the future state of federal evidence use and, therefore, will be reported in our next report (i.e., Report 3: Future State of Evidence Use by Federal Policy Stakeholders). In this report, we focus on the recent past and current challenges reported about federal data use experiences. These alone, help inform the needs and reported pitfalls across data systems, but still map to parts of the proposed NSDS model as suggestions regarding promising features that could address existing challenges and influence the federal data ecosystem.

There were initially 48 challenge themes identified across the 12 KIIs, however, some were duplicates or similar enough that the coding team discussed and agreed upon consolidating similar themes to generate a set of mutually exclusive themes. In Appendix E, a table shows the matrix of the coded themes aligned with a social or technical system (columns, each with two components) and part of the proposed NSDS model (rows). This table only shows the coded challenge themes in this report. Report 3 will show a fuller version of this table with proposed solutions, opportunities, and suggestions coded themes as signals for the future state of federal evidence use.

There are many interesting patterns that can be observed from the challenges alone. First, after deduplicating the redundant themes, there were **21 unique challenge themes** that emerged from the KIIs (i.e., narrowed down from the original 48 challenge themes). This indicates that many of the informants experienced similar challenges and there were several challenge themes that were common regardless of the informant's role, organization type, or expertise. Another early observation noted in Table 4 in Appendix E is that all the themes that did not map to the NSDS model are external threats that may impede the use of federal data. These challenges are likely too complex or high-level and perhaps outside the current scope of this project to be addressed within the framework we are developing for the future NSDS.

Additionally, **all social system challenges at the organizational-level[11] are external barriers** that do not map directly to the NSDS. This means that these are challenges related to organizational climate, behavior, and culture that the NSDS cannot (and, likely, will never be able to) address because those challenges are deeply rooted in historical and institutional norms that are difficult to overcome. These organizational-related challenges include:

- Incongruence Between Policy Needs and Evidence Timing
- Resource vs. Demand Imbalance
- Cultural Resistance to Sharing and Using Data
- Shifting Priorities and Unstable Trust
- Evidence Act Implementation Limitations

The challenge themes clustered in interesting ways across each of the proposed NSDS model layers. Challenges related to the **Discovery and Navigation** layer were Social-Individuals or Technical-Technology challenges. This means these challenges were mostly related to building a technical infrastructure to support efficient data discovery and navigation to help individuals better identify possibly relevant data and subject matter expertise for their policy needs. Challenges related to the **Shared Services and Resources** layer were Social-Individuals or Technical-Tasks challenges. This means these challenges were mostly related to building the capacity and networks of individuals to support better question formulation and through complex analyses. Lastly, challenges related to the **Access and Linkage Infrastructure** layer were Technical-Tasks or Technical-Technology challenges. This means these challenges were mostly related to building the technical processes and systems to facilitate better data quality, access, sharing, interoperability, and integration.

3.3. Detailed Summary of Challenge Themes with Descriptions

The following are detailed descriptions of each of the 21 challenge themes that emerged from KIIs with I3P members. These themes represent the attitudes, knowledge, and experiences of the informants who were interviewed. Embedded quotations reflect direct statements that were made by deidentified I3P members.

3.3.1. Data Discovery and Navigation Challenges

3.3.1.1. Identifying and Mapping Relevant Stakeholders

Direct data users struggle to identify who has influence and expertise on specific topics, often relying on "usual suspects" from elite institutions rather than discovering diverse voices. One interviewee noted the difficulty of "figuring out who to talk to, who had sway and influence on a topic," while organizational charts do not reveal actual influence patterns or interpersonal relationships that drive policy decisions.

3.3.1.2. Congressional Data Literacy Gap

Most congressional staff lack basic data skills required for their roles, with many being young professionals with limited analytical background who are unaware of existing resources. Staff often want final products rather than guidance on methodology and face immediate turnaround expectations that conflict with the time needed for proper data analysis.

3.3.1.3. Limited Information Discovery and Synthesis

Direct data users face significant challenges navigating existing data systems, with platforms like data.gov described as inadequate for actual use. The need to pull together diverse information sources for policy arguments is complicated by time constraints and heavy reliance on contractors for technical analysis.

[11] This system code describes themes related to the policies, priorities, procedures, resources, and culture that establish the conditions under which policy work is done.

3.3.1.4. Difficulty Navigating Existing Data Sources

Federal data systems are poorly designed for user navigation, requiring extensive technical knowledge and institutional relationships to access relevant information. Multiple databases with different search capabilities and training requirements create fragmented access, while inconsistent tagging and lack of standardized terminology prevent users from easily answering basic questions about agency activities.

3.3.2. Shared Services and Resources Challenges

3.3.2.1. Questions Limited by Policymaker Data Awareness and Literacy

People do not ask questions they assume cannot be answered, leading to self-censoring of potentially valuable analytical inquiries. Limited knowledge of data capabilities among potential users means that policy questions are often constrained by what decision-makers believe is possible rather than what data could actually reveal.

3.3.2.2. Need to Increase Staff Capacity and Skills

There is a critical need for both technical skills and domain knowledge, with the challenge being easier to find technical programmers than people who understand policy nuances. Multiple interviewees emphasized that analytical skills are often learned through mentorship rather than formal training, making capacity building difficult to scale across government.

3.3.2.3. Question Formulation Challenges

Policymakers frequently start with broad, unfocused questions rather than specific, empirically testable ones, with many beginning inquiries with "how can we" rather than formulating researchable hypotheses. The disconnect between idea-generating policy brainstorming and empirically analyzable questions requires significant refinement work to translate policy goals into specific, data-driven analyses.

3.3.2.4. Complexity of Accessing Data

Data access involves lengthy, complex approval processes that can "easily take a year, if not more," with each data source requiring separate agreements and coordination between different Information Technology (IT) systems. Legal and administrative barriers create significant obstacles, while risk-averse federal employees with competing responsibilities struggle to dedicate time to complex data negotiations.

3.3.3. Access and Linkage Infrastructure Challenges

3.3.3.1. Complexity of Aggregating and Coordinating External Data Sources

Combining datasets from multiple agencies requires extensive coordination between different data managers, IT systems, and legal frameworks. Each data linkage requires separate, complex agreements, while agencies often have different rules about data commingling and may only provide customized extracts rather than full integration capabilities.

3.3.3.2. Interoperability and Metadata Challenges

Administrative and statistical data serve different purposes, leading to user confusion about which sources to use, while insufficient documentation exists about data meaning and appropriate applications. Different agencies use inconsistent definitions for the same categories, and metadata gaps prevent users from understanding data limitations and quality issues.

3.3.3.3. Onerous Processes to Access and Share Data

Complex data use agreements, multiple approval layers including cyber and technical review boards, and extensive legal consultations create bureaucratic barriers that discourage data sharing. Lawyers are often "very anxious" about data sharing even for aggregate information,

while precedent-setting becomes crucial for overcoming basic risk aversion in government agencies.

3.3.3.4. Data Infrastructure and Integration Issues

Legacy systems designed for (payment) transactions rather than analytical purposes create fundamental mismatches with research needs. Multiple incompatible systems from different vendors, limited secure data access capabilities, and system fragmentation prevent effective data integration across agencies and programs.

3.3.3.5. Poorly Defined Data Governance and Usability Responsibilities

There is an absence of clear roles and responsibilities for data governance, with agencies lacking standardized frameworks and metadata management systems. Chief Data Officers often have "no staff" and lose "turf battles with Chief Information Officers," while organizational structures do not support systematic data sharing and quality assurance.

3.3.4. External Cultural and Environmental Factors

3.3.4.1. Incongruence Between Policy Needs and Evidence Timing

Political decision-makers operate on short timelines and may not wait for evidence, while sophisticated data analysis requires months or years to complete properly. The disconnect between when surveys are designed and when policy questions emerge, combined with administrative data limitations, creates persistent timing mismatches between evidence availability and policy needs.

3.3.4.2. Resource vs. Demand Imbalance

Small agencies have limited funding compared to larger organizations, while increasing recognition of the value of data leads to growing requests that exceed analytical capacity. Staffing challenges mean agencies often cannot meet demand for data analysis, with some having lost the majority of their quantitative analysts while facing pressure for quick policy responses.

3.3.4.3. Cultural Resistance to Sharing and Using Data

Agencies avoid sharing data without established precedents due to legal risk aversion, while there is widespread resistance to evaluation terminology and evidence-based approaches. The fundamental challenge is that "for most people, data are just another opinion," indicating a broader cultural deficit around evidence-based decision making that extends beyond technical solutions.

3.3.4.4. Shifting Priorities and Unstable Trust

Administration changes and annual budget appropriation cycles disrupt multi-year evaluation initiatives and halt strategic planning efforts, while inter-agency councils can be suspended during political transitions. The instability creates challenges for long-term data projects and relationship building, as personnel changes and shifting political priorities undermine consistent collaboration frameworks.

3.3.4.5. Evidence Act Implementation Limitations

The Evidence Act is described as "in real trouble" due to lack of funding for implementation requirements and insufficient technical capacity among advocates who "were not technologists." Evaluation officers lack leadership access and influence, while legal frameworks remain complex obstacles to data sharing despite legislative intent to improve evidence-based policymaking.

3.3.4.6. Influence of Organizational Culture on Individuals

Even at science agencies, there is a "far cry from doing evidence-based policymaking," while terminology matters significantly in gaining acceptance for analytical work. Cultural factors influence whether people are comfortable asking questions and using data for decisions, with organizational environments either supporting or discouraging evidence-based approaches regardless of individual preferences.

3.3.4.7. Staffing and Institutional Knowledge Loss

Agencies face retirement waves of experienced personnel, with "a lot of people trying to retire" taking institutional knowledge with them. Critical skills gaps exist where expertise requires both technical abilities and domain knowledge, while staff reductions have eliminated key quantitative analysts, creating major capacity gaps at agencies trying to maintain analytical functions.

3.3.4.8. Dependency and Overrepresentation of Usual Contractors

Government work relies heavily on contractors for analytical capacity, with relationships spanning decades being essential for quality work. There is overrepresentation of elite institutions and well-known organizations in contractor selection, while the contracting structure prevents retention of institutional knowledge and creates succession planning challenges when senior expertise is difficult to replace.

3.4. Three Universal Challenges Shared Across I3P Members

The greatest overall challenge articulated by every informant is the **mismatch between policy needs and the timing of available data**. The short timeframe for meeting policy demands was contextualized as resulting in “not enough time”, “focusing on quick wins”, “Congressional timeline pressures”, or “limiting possible analyses”. Because this factor is an external barrier to NSDS implementation and use, there is little the future NSDS can do to change the timing and urgency of policy demands. However, the future NSDS can be a useful tool in expediting timely and relevant data to respond to anticipated policy demands.

While there are some individual-level challenges within the social system that are external barriers (e.g., Staffing and Institutional Knowledge Loss, Dependency and Overrepresentation of Usual Contractors) that threaten NSDS implementation and use, the greatest challenges related to individuals were weaknesses that may be addressed by proposed features within the Discovery and Navigation and Shared Services and Resources layers of the proposed NSDS model. Most notably, almost all informants expressed the **need to increase staff capacity and skills** to conduct analyses with greater volume, quality, sophistication, and efficiency. This need is something that the future NSDS could address through the Shared Services and Resources layer of the proposal model, specifically with the capacity building center and toolkits.

Another universal theme across nearly all informants was an emphasis on the **arduous process of obtaining data sharing agreements** across agencies and the **inability to easily link data** across systems. These are both process (tasks) and systems (technology) issues that severely impede the already pressing timelines to meet urgent policy demands. Through the future NSDS Access and Linkage Infrastructure layer, standing agreements and a more robust data infrastructure designed to support complex linkages across systems would completely revolutionize how data are shared and used for federal policy purposes.

3.5. Illustrative Vignettes of Real Challenges

In conducting the KIIs, it became clear that the informants were sharing various experiences from different roles across several organizations and time periods of their extensive careers. This meant that many of the experiences they shared were partial stories, fragments, or

features of their path to and through using federal data, but there were very few stories that covered the whole process from beginning to end and, inherently, none that could provide the perspectives or experiences of others involved in each instance. However, what we could do was get a glimpse into one clear story for each informant that demonstrated the complexity and nuances of a policy product or need, challenges experienced in trying to produce it, and the current strategies undertaken to achieve the intended goal.

We refer to these stories as "illustrative vignettes", which are user stories that describe a specific experience of accessing and/or using federal data for policy purposes. The table in Appendix F provides illustrative vignettes we highlighted, one for each of the I3P informants. Some of the vignettes were from previous roles or organizations. Although the table includes anonymized, generic identifiers to classify the informant and their organization, that information is presented according to the role and organization type that aligns with the vignette and may not characterize their current role or organization.

Through the process of developing these illustrative vignettes, we came to realize that these stories begin to piece together parts of a larger journey map of how different stakeholders work with federal data at different stages of their career. Through their innovative strategies to navigate these real, present challenges, our I3P informants give us a glimpse into the efficiencies and mechanisms needed to meet policy demands.

3.6. Identifying the "Unmet Needs" of the Analyst Persona

The need for an additional journey mapping method emerged early in the study design process. We became aware of a need for technical assistance expressed by analysts staffing a federal policy advisory Board. Although our I3P KIIs did not include any (current) direct data users of federal data, we held separate conversations with analysts assigned to support this Board to better understand their challenges and unmet needs.

The Board and their analysts invited us to offer suggestions for re-formulating nebulous, complex policy problems as well defined problems that would lend themselves to systematic and rigorous evidence based analysis and then identify and determine where to find and how to access data that could serve as input for building evidence to offer policy recommendations to Congress. We seized the opportunity to not only describe the analyst's user journey as it happens in real-time, but also to identify subject matter experts (SME) who could offer guidance. In doing so, we framed our approach as a use case[12] that demonstrates a number of potential ways that the NSDS Concierge Service could use SMEs to address some of the unmet needs of analysts, and the policy stakeholders they serve.

This use case essentially produced a "pseudo-prototype" but is not intended to be an example of the rapid, low-fidelity prototype development and testing that can be part of the HCD process. The details of this use case are described in a confidential report provided to the host agency. The de-identified analyst persona will contribute to a comprehensive journey map constructed from this use case along with perspectives from the KIIs will be presented in our next report.

3.7. Mapping the Journey to and through the Federal Data Ecosystem

Through the KIIs and use case, we heard several accounts from different perspectives about the path that direct data users take when conducting research inquiry on any issue or for any purpose. Mapping out this journey to and through the research inquiry helps elucidate the

[12] For the purpose of this study, a use case is defined as an opportunity where a team can "search and test ideas regarding their needs, challenges, and desires" (source). This method complements the retrospective approach to surfacing the unmet needs of potential users of the future NSDS.

fundamental steps in the process of accessing and using data for any type of inquiry, and the tasks within each step. Graphically, this framework is called a **user journey map** using the generic inquiry lifecycle as an anchor and organizing framework. A journey map provides a visual representation of the route that direct data users must take to fulfill the requests made of them, and the challenges they may encounter along the way.

Given the KIIs and use case offer a rich understanding of the challenges experienced by data users, we can not only map the steps in the journey and describe the nature of each step but also map reported challenges to the journey to gain a deeper appreciation of how the 12 challenge themes (that align with different levels of the proposed NSDS model as shown in Table 4 in Appendix E) map specifically to the policy inquiry process.

The following journey map shown in Figure 2 was constructed from the use case about unmet needs described in the previous section (i.e., section 3.6), which were also confirmed through KIIs with the I3P members. This figure provides an outline of the 3 phases of the inquiry lifecycle (Planning & Design, Implementation, and Dissemination) and 6 steps across those phases: Clarify Request, Scope Requirements, Acquire Data, Prepare Data, Fulfill Request, and Communicate Insights. Although the steps may seem like a unidirectional and linear model, the process is iterative and has informal feedback loops that may cause the user to backtrack or toggle between multiple steps. The figure below also briefly describes direct data user tasks required within each step along the journey map.

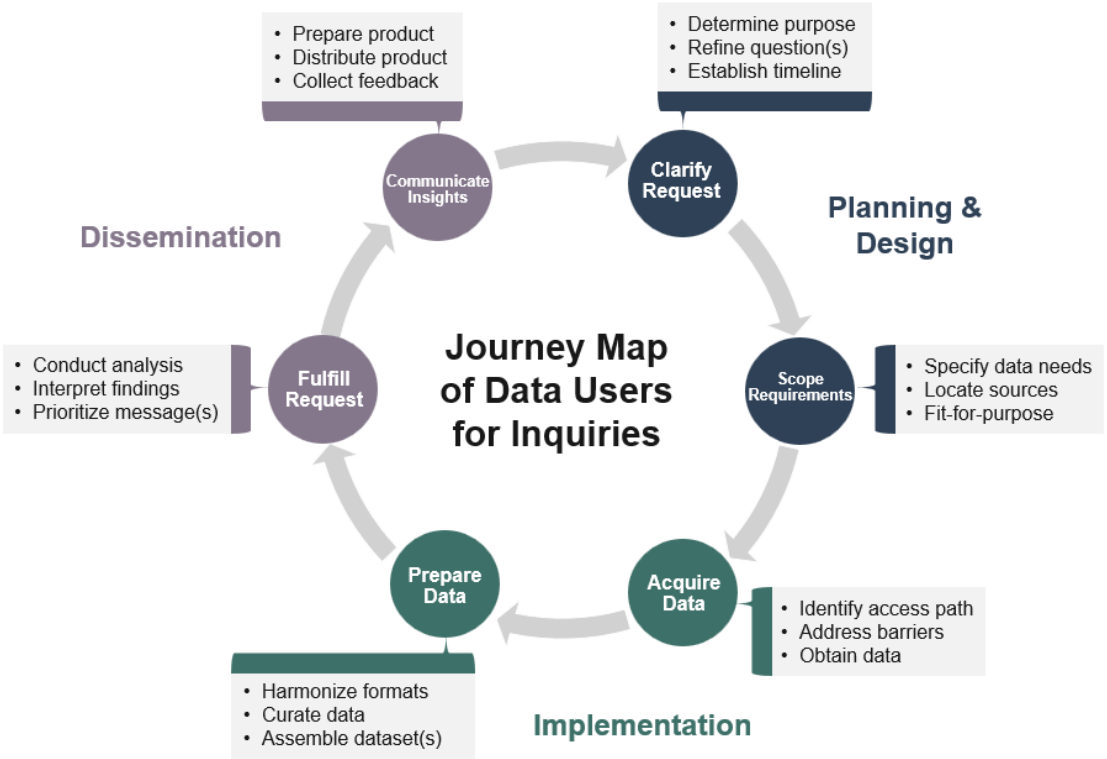

Figure 2. Overview of the Direct Data User Journey Map for Policy Inquiry

In Appendix G, Figure 3 displays the same journey map but in a linear fashion to make some clearer connections to the challenges expressed by direct data users in the use case and throughout the 12 KIIs. The analysts from the use case articulated myriad challenges that are often riddled throughout the entire policy inquiry process. In Figure 3, these direct data user challenges are summarized under each step as “Task-Level Challenges in the Policy Context”.

Furthermore, the 12 challenge themes that emerged from KIIs that mapped to proposed NSDS layers were also mapped to the journey map. Because the I3P KIIs provide a more senior-level, strategic perspective on the challenges, the challenges experienced and expressed by the analysts relate directly to the tasks in the journey map while the challenge themes speak more to system-level issues. While the journey map describes any inquiry process, the challenges contextualize the nuances and pain points specific to inquiry done for policy purposes.

4. Conclusions

Across the 12 KIIs and a use case, participants shared their experiences interacting with federal data for different policy purposes in varying roles and organizations. While no one informant provided an end-to-end description of the use of federal data across an entire policy lifecycle, together, their perspectives surfaced unmet needs. Through their individual stories, which represent many snippets of federal data use across various contexts, we are able to begin piecing them together to comprehend the larger policy landscape that highlights the varied ways in which different actors use federal data, the factors and the circumstances that influence their inquiries a framework for how a future NSDS could help meet those needs. One major takeaway is the need to be more proactive about anticipating policy questions. This anticipation can be achieved by continuously surveilling agency strategic plans (and corresponding learning agendas) and emerging issues. Skills, information, and systems can then be better designed to address policy questions efficiently and effectively.

In our next report, we will build on the findings reported here and include proposed solutions, opportunities, and suggestions that emerged from our interviews. Already, we are seeing that the user journey is heavily influenced by their role and needs (i.e., their persona type and career stage), the policy purpose (i.e., policy question specificity and location along the policy lifecycle), and the proximity and context of their needs to the data. These elements will be seamlessly woven together into a single, comprehensive journey map and SWOT analysis that articulates the path of user experience and the internal and external factors that will contribute to successes and challenges for future NSDS implementation and use. While there are barriers outside the purview of the NSDS that are difficult to overcome, the illustrative vignettes are evidence that motivation drives innovation to get tasks done. Several current strategies employed by I3P informants and their colleagues were a result of not accepting the *status quo* of the federal data ecosystem's challenges. Sustaining and enhancing these solutions requires money, mandates, and motivation. Although the future NSDS cannot directly affect any of the organizational barriers, it can be a solution that offers cost reductions, standards and precedents, and drive motivation to change attitudes and behaviors of federal data users.

5. Next Steps and Timeline

Our analysis reveals that, while there is strong appetite for improved data access and integration, current agencies and the federal statistical systems fail to meet the speed, usability, and coordination requirements of policymakers. The most significant barriers are not technical but organizational—involving data sharing agreements, institutional silos, and misalignment between analytical capabilities and policy timelines. Ultimately, success will require a human-centered approach that prioritizes rapid response, question refinement, and relationship building once the NSDS 1.0 is built as a primarily technical solution.[13]

[13] Pahlka, J. (2023) *Recoding America: Why government is failing in the digital age and how we can do better*. Metropolitan Books.

6. Appendices

Appendix A: I3P Panel Member Demographics

Table 2. I3P Panel Member Demographics*

Policy Role	Organization Type	Years of Policy Experience	Years in Government
Policymaker	Regulatory	17	19
Policymaker	Public Good (State)	19	13
Policymaker	Public Good	18	13
Policy Advisor	Science	15	12
Policy Advisor	Science	5	17
Policy Advisor	Public Good	10	18
Policy Evaluator	Public Good	13	13
Policy Evaluator	Regulatory	16	16
Policy Evaluator	Science	16	6
Policy Influencer	Non-Governmental (Nonprofit)	10	0
Policy Influencer	Non-Governmental (Nonprofit)	19	12
Policy Influencer	Non-Governmental (Academic)	5	5

Appendix B: I3P Interview Guide

We will focus our conversation on the example you selected to describe your user experience seeking federal data to address a policy issue of national importance.

Individual, Culture, & Context Characteristics

1. Briefly describe your role(s) and responsibilities relative to policy inquiry at your agency?
    a. The Evidence Act? How do you describe the role of data in the policy process? Similar or different from those you work with and for?
2. What is the policy issue or policy question that you wanted to use federal data to address?
3. Where in the policy lifecycle were you involved?
    a. Show diagram of the policy lifecycle. If there is more than one place, ask the respondent to select the area where most of their user experience occurs.
4. At what level in the agency or your organization did the question come from?
    a. Direct use or indirect “bridge” support.

Reflection on the Current User Experience

5. What did you need from the data? Were there specific points you needed to support or discover?
6. Where did you go to get the data?
    a. Describe the type of products you were looking for How did you envision using these data at the outset?
7. Tell me about the steps you took to find what you were looking for.
    a. Problems, pain points, feelings, high points, tools, resources?
8. Who, if anyone, helped you along the way, and how?
    a. What was their role? What did they do to help you? Within your agency or elsewhere?
9. What did you do with the data?
    a. Build evidence? Problems, pain points, feelings, high points.
10. What, if anything, happened as a result of your experience? Did the policy question or issue change during the process?
    a. How? Why? Who? Were the data used for policy purposes? Why or why not?

Improving the Future User Experience

11. Let’s look at a diagram that describes the features envisioned for a future NSDS. What do you see here, if anything that might have helped you during the scenario we discussed?
12. What new policy issues, if any, are on the horizon for you?
    a. How might you approach improving your use of federal data for policy purposes?

Appendix C: Prompts for Claude.ai Summaries and Verification

Summary of Raw Interview Transcript

"Please provide a structured summary of the following interview transcript. Use these as the headers for the summary: key points (including interview overview, key background context and current work examples); key challenges; proposed solutions and NSDS applications; user categories and needs; opportunities for inter-agency coordination; critical insights and recommendations for NSDS design; and future engagement opportunities."

Fidelity Verification

"The first document is a summary that you generated of the second attachment, which is the raw transcript from a key informant interview. Please show me evidence of how you created your summary, including direct quotes that corroborate your interpretation and any other examples that help ensure the summary is free of hallucinations."

Summary of Summaries

"Please provide a structured summary of the following 12 interview transcripts, noting both what was common across all 12 interviews and unique elements. Use these as the headers for the overall summary: key points (including interview overview, key background context and current work examples); key challenges; proposed solutions and NSDS applications; user categories and needs; opportunities for inter-agency coordination; critical insights and recommendations for NSDS design; and future engagement opportunities."

Individual Phases and User Categories of the Policy Process

"Create a structured summary of the attached interview transcript to extract what the informant said about their experience using federal data. Organize their experience across clear steps in a process of inquiry and, where possible, include any substeps that are mentioned and the actors that support them along the way."

Summary of Phases and User Categories Across All Interviews

"Create a structured summary of the attached interview summaries to extract what the informants said about their experiences using federal data. Organize their experiences across clear steps in the process of inquiry and, where possible, include any substeps that are mentioned and the actors that support them along the way. Synthesize and summarize any other content."

Detailed Descriptions of Each Challenge and Solution Theme

"Using only the summary documents attached here, please write 2-3 sentences summarizing the high-level description for each of these themes that emerged from these documents. Do not identify any of the interviewees in the descriptions."

Appendix D: Codebook Used for Qualitative Coding

Table 3. Codebook Used for Qualitative Coding

Code Type	Codes	Descriptions
Systems	Social - Organizations	Policies, priorities, procedures, resources, culture--establish the conditions under which policy work is done.
	Social - Individual	Knowledge, networks, skills--capacity to do policy work.
	Technical - Tasks	Workflows, operations, job--the steps required to engage in policy work.
	Technical - Technology	Infrastructure, tools, resources--aids to improve the efficiency and/or effectiveness of policy work.
SWOT Analysis	Strength	Internal positive attributes, resources, capabilities, or advantages that give the organization/entity a competitive edge or enable successful performance.
	Weakness	Internal negative factors, limitations, or deficiencies that hinder performance or place the organization at a disadvantage relative to competitors.
	Opportunity	External positive factors, trends, or circumstances in the environment that the organization could potentially exploit or leverage for growth and competitive advantage.
	Threat	External negative factors, challenges, or risks in the environment that could potentially harm the organization's performance or competitive position.
NSDS Model Mapping	Discovery and Navigation	Front door and concierge service providing visitors with resources and services or directing them to the right resources and services to meet their needs for data and data infrastructure.
	Shared Services and Resources	Four main resources and services available to everyone who accesses the NSDS: toolkits (tailored to specific needs), capacity building center (training), communities of practice, and data usage platform.
	Access and Linkage Infrastructure	Restricted access to a secure computing environment using state of the art privacy-preserving technologies and data linkages.
	External Cultural and	External environmental barriers or facilitators that would negatively or positively influence the

	Environmental Factors	effectiveness and efficacy of NSDS implementation or use.

Appendix E: Matrix of Challenge Themes by System and Mapped to NSDS Model

Table 4. Matrix of Challenge Themes Organized by System and Mapped to NSDS Model

	Social - Organizations	Social - Individuals	Technical - Tasks	Technical - Technology
Discovery and Navigation	• No Challenge Themes	• Identifying and Mapping Relevant Stakeholders* • Congressional Data Literacy Gap	• No Challenge Themes	• Limited Information Discovery and Synthesis • Difficulty Navigating Existing Data Sources*
Shared Services and Resources	• No Challenge Themes	• Questions Limited by Policymaker Data Awareness and Literacy* • Need to Increase Staff Capacity and Skills*	• Question Formulation Challenges • Complexity of Accessing Data*	• No Challenge Themes
Access and Linkage Infrastructure	• No Challenge Themes	• No Challenge Themes	• Complexity of Aggregating and Coordinating External Data Sources* • Interoperability and Metadata Challenges • Onerous Processes to Access and Share Data* • Poorly Defined Data Governance and Usability Responsibilities*	• Data Infrastructure and Integration Issues* •
External Cultural and Environmental Factors	• Incongruence Between Policy Needs and Evidence Timing* • Resource vs. Demand Imbalance* • Aggregation and Coordination of External Data Sources* • Shifting Priorities and Unstable Trust* • Evidence Act Implementation Limitations*	• Influence of Organizational Culture on Individuals • Staffing and Institutional Knowledge Loss* • Dependency and Overrepresentation of Usual Contractors*	• No Challenge Themes	• No Challenge Themes

Color Coding: Strengths; Weaknesses; Opportunities; Threats
* composite themes that represent several individual but similar themes

Appendix F: I3P Illustrative Vignettes

Table 5. Highlights from I3P Interview Illustrative Vignettes

Policy Role	Organization Type	Policy Process[14]	Policy Artifacts	Challenges	Current Strategies
Policymakers[15]					
Agency Senior Leader	Regulatory	Policy Implementation	Agency Capacity Assessment	Data silos	Coordinated across the agency to move toward a more robust data infrastructure
Agency Senior Leader	Public Good (State)	Policy Formulation	Policy Issue Presentation	Inaccessible language for the intended audience	Collaborated with communication specialists to develop accessible and relevant messaging to reach policymakers
Agency Senior Leader	Public Good	Policy Evaluation	Return on Investment Reports	Limited staff capacity and capabilities	Had senior level supervisors manage both the data and policy analysts
Policy Advisors[16]					
Policy Intermediary	Science	Policy Formulation	Policy Brief	Limited staff and availability of information to respond to issues urgently	Used an advisory group that identified and curated a list of experts to support staff in crafting timely and relevant briefs
Policy Intermediary	Science	Problem Identification	Convening(s) and Issue Briefs	Lack of prioritized policy issue opportunities	Facilitated group consensus-building to prioritize policy issues
Policy Intermediary	Public Good	Policy Implementation	Data Issue Presentation	Lack of interagency linkages	Piloted cross-agency data linkage projects to demonstrate feasibility and promote utility
Policy Evaluators[14]					

[14] Public policy process or lifecycle phases include problem identification, policy formulation, policy implementation, and policy evaluation (Dunn, 2018; Peters, 2018; Weible and Sabatier, 2018).
[15] Individuals at the highest levels of agency decision making on substantive policy topics of Congressional interest (e.g., Advisory Board and Committee members, Senior Leadership, Chief Data Officer, Chief Evaluation Officer and Chief Statistical Official).
[16] Senior-level government employees, who serve as intermediaries balancing the evidence needs of agency leadership directly shape the sociotechnical systems at the agency level for policy purposes. Agency and government-wide staff with responsibilities for determining and/or providing evaluative evidence on the efficiency and effectiveness of agency Federal policies and programs.

Policy Intermediary	Public Good	Policy Evaluation	Evidence Briefs	Sophistication of analyses and ability to respond to policy issues rapidly	Identified anticipated evidence gaps and strategically planned better ways to generate relevant data to address gaps
Policy Intermediary	Public Good	Policy Evaluation	Program Evaluation Report	Linking data from the statistical system with program administrative and evaluation data	Involved program staff in the planning and design of the program evaluation
Policy Intermediary	Science	Policy Implementation	Dashboards	Incongruency between administrative and evaluation data	Collaborated with various stakeholders to meet information needs at the right level of analysis
Policy Influencers[17]					
Policy Communicator	Non-Governmental (Nonprofit)	Problem Identification	Fact Sheets	Issue data are in multiple locations and are often difficult to find or access	Made data readily available into useful and responsive fact sheets
Policy Communicator	Non-Governmental (Nonprofit)	Policy Formulation	Examples of Effective Policy Implementation	Making evidence relatable and accessible in timely manner	Funded implementation projects and created portfolio exemplars to communicate what works
Policy Communicator	Non-Governmental (Academic)	Problem Identification	Data Governance Presentation	Insufficient protocols and infrastructure for data governance	Raised awareness and a possible solution to standardize data governance across government

[17] Individuals outside the agency who engage in policy education and other activities related to substantive agency policy direction of Congressional interest.

Appendix G: Direct Data User Journey Map to and through the Federal Data Ecosystem

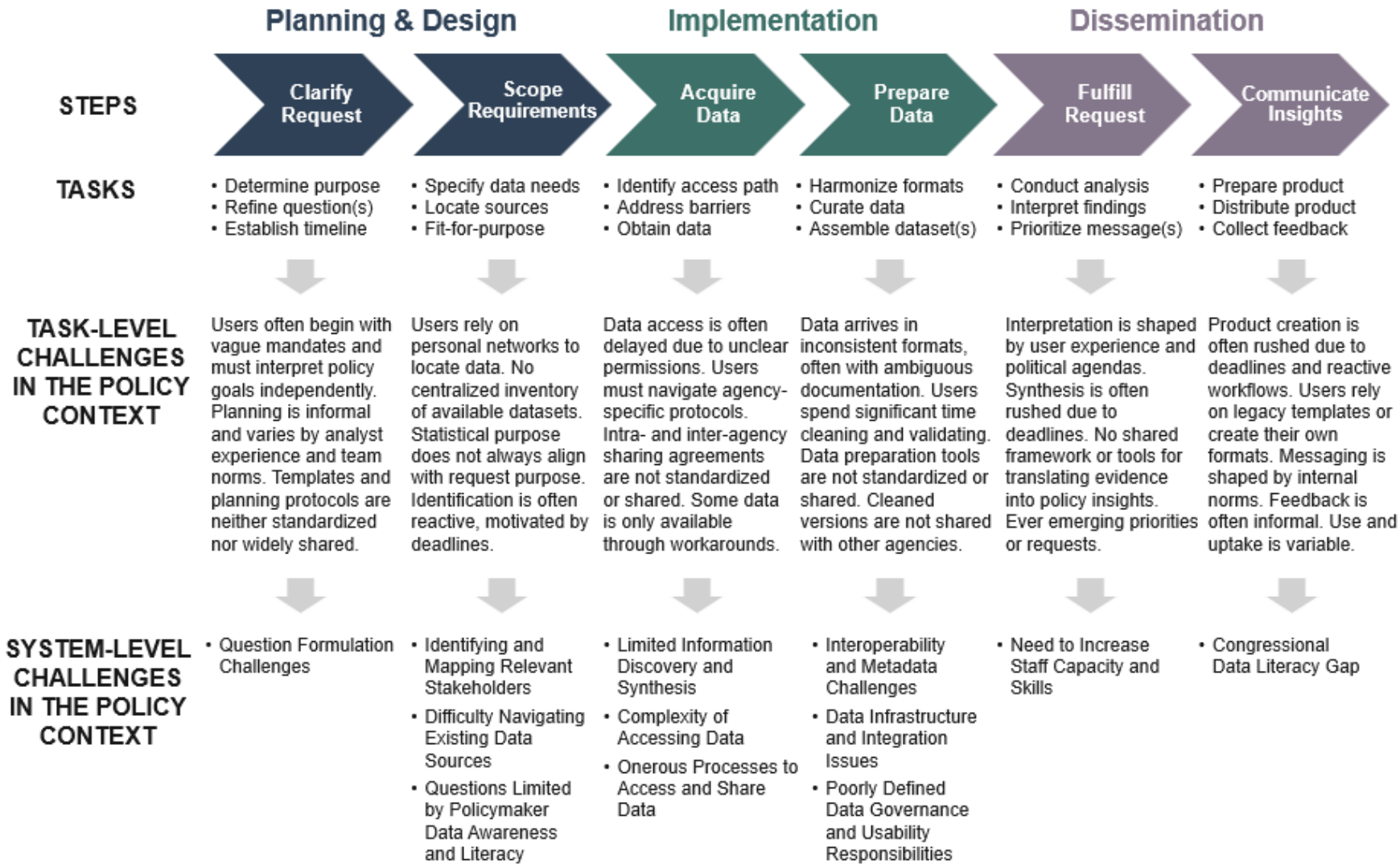

Figure 3. Direct Data User Journey Map to and through the Federal Data Ecosystem