



# Technical Infrastructure Focus Area

**Co-Chairs, Amy O'Hara and David Park, ACDEB Members**

July 2022

*Members:* Otis Brown, Leonard Burman, Barry Johnson, Ted Kaouk, Elisabeth Kovacs, Mayank Varia, Christina Yancey

# Draft Recommendations

---

## Focus Area Discussion

### NSDS Functional Capability: Coordination

**Recommendation 1 – Data Accessibility:** The NSDS will provide technology so that users with various tiers of access can safely, easily, and efficiently analyze data assets hosted by affiliated organizations, including Federal, State, local, tribal governments, non-profit and other organizations. To support discovery of data assets for evidence building, the NSDS will provide a technological process to support access to discoverable metadata, request data access, track the approval process, and document the outcome.

**Recommendation 2 – Technical Assistance:** The NSDS will also provide technical support for a concierge service to aid data contributors in deciding which access controls to apply to their datasets and to aid analysts in selecting datasets and obtaining access for approved uses. This support should enable both automated (bots) and in person assistance and provide usage metrics.

# Draft Recommendations

---

## Focus Area Discussion

### NSDS Functional Capability: Coordination

**Recommendation 3 – Data Integrity:** The NSDS will provide tools and support to users in conducting secure, accurate, and scalable analytical and evidence-building activities such as tabulations, dashboards, regressions, record linking, and machine learning. The NSDS will support the development and deployment of data access protocols that offer alternatives to the standard direct data access models, such as synthetic datasets with access to a validation or verification server to check the integrity of results.

**Recommendation 4 – Data Auditability:** To ensure auditability and transparency, the NSDS will provide tools to track the provenance of all datasets and all evidence-building analyses performed over these datasets. This will include tools for accessing and reporting on the quality of data assets, aligned with recommendations from the Federal Committee on Statistical Methodology (FCSM). To promote open and reproducible science, the NSDS will provide public attestation of the datasets and procedures used in a data analysis, upon request.

# Draft Recommendations

---

## Focus Area Discussion

### NSDS Functional Capability: Coordination

**Recommendation 5 – Disclosure Limitation:** NSDS will facilitate the development and application of statistical disclosure limitation methods. This activity will include infrastructure to host tools and trainings included in the Data Protection Toolkit developed by the FCSM Committee on Data Access and Confidentiality and an active research program to develop new approaches and training. To ensure that methods applied are developed using realistic risk models, the NSDS will provide users with infrastructure and support for conducting privacy risk assessments.

**Recommendation 6 – Data Linkage:** The NSDS will advance privacy-preserving record linkages to reduce risks of unauthorized access on existing data joins and to overcome barriers that prevent linkages across sources. The NSDS will provide tools and support to users in conducting scalable, privacy-preserving record linkages, facilitating data preparation and output review. The data concierge service will also coordinate with state, local, and tribal government officials seeking linkage services.

# Draft Recommendations

---

## Focus Area Discussion

### NSDS Functional Capability: Communication

**Recommendation 7 – Knowledge sharing:** NSDS should provide guidance to users and the general public about the evolving state of the art in solutions for evidence-building in a privacy-preserving, publicly auditable, and fair and equitable manner.

Additionally, NSDS should provide web-based infrastructure to foster communities of practice to support cross-discipline work and knowledge sharing, in coordination with federal funding agencies such as NSF, NIH, and DARPA that invest in foundational research in these areas.

# Draft Recommendations

---

## Focus Area Discussion

### NSDS Functional Capability: Research and Development

**Recommendation 8 – Innovation Sandbox:** NSDS should feature a sandbox for testing new and innovative technologies and software for multiple data access tiers, data protection protocols, and data analysis. The NSDS should be a neutral ground between agencies where secure testing of new data linkages, privacy-preserving technologies, and model approaches can occur.

**Recommendation 9 – Privacy Preserving Technologies:** NSDS will promote use of privacy technologies that support working with data in situ, working with the research community to develop efficient, scalable tools for users from all levels of government (including open competitions). NSDS will help build capacity through training and technical support for users deploying these technologies.

# Full Committee Discussion

---

## Discussion Questions

- Are there any showstoppers or things to be clarified in the text itself—noting that the supporting findings are still under development? Any insights to highlight from the information gathering or findings of other subcommittees?
- Do these recommendations capture the appropriate spectrum of technical infrastructure needed to support the Committee’s vision for the NSDS? Why or why not? Are items missing? Is there too little/too much emphasis on any of the areas?
- There are overlaps with these recommendations and those of the other subcommittees (for example, technical assistance and communications), what is the best way to integrate these items? Are there things the subcommittees can do now to help with synthesis/integration?