

Safe Data Technologies Project

Project Funding & Collaborators



**ALFRED P. SLOAN
FOUNDATION**



Project Goal

*This body of work aims to **safely expand access to confidential data** that **advances evidence-based policy making** by creating new ways for researchers to use administrative data while protecting privacy.*

Project Framework

*Our work is at the intersection of data privacy and public policy. We are **implementing practical privacy-preserving technologies and tools** (e.g., synthetic data generation) and **exploring the feasibility of state-of-the-art methodologies** (e.g., formal privacy) to provide better data access.*



THE PROMISE OF EVIDENCE-BASED POLICYMAKING

Report of the Commission on Evidence-Based Policymaking



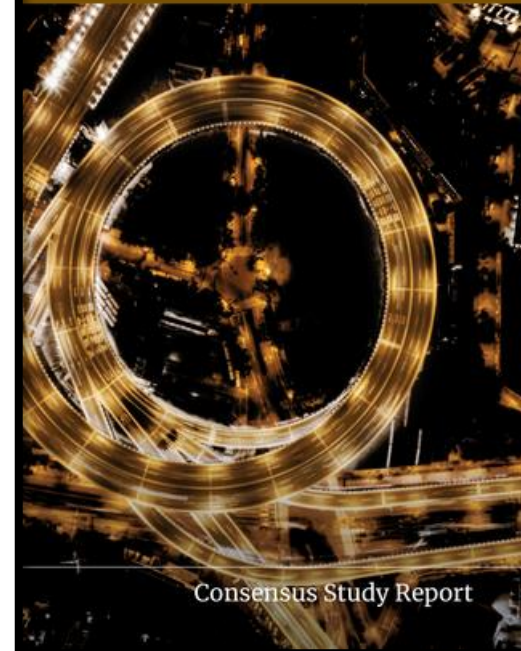
Advisory Committee on Data for Evidence Building: Year 2 Report

October 14, 2022



NATIONAL ACADEMIES
Sciences
Engineering
Medicine

Toward a 21st Century
National Data Infrastructure:
Managing Privacy and Confidentiality
Risks with Blended Data



Consensus Study Report



AMERICA'S DATAHUB
C O N S O R T I U M

Tiered Access for Administrative Tax Data

Our goal is to enable more researchers to safely access confidential tax data.

- I. **Basic Access:** Summary tables on the Statistics of Income Division website.
- II. **Public Use File (PUF):** Researchers will have access to the synthetic PUF.
- III. **Validation Server Access:** Researchers are trusted to access the validation server, where they submit statistical analyses that they have tested and debugged on the synthetic PUF. Researchers at this tier will have to undergo an application process.
- IV. **Full Access:** Researchers who obtain clearance and therefore have access to the unaltered, confidential data, but will be still be limited on what information can be released.

Safe Data Technologies Project: Subteams

*Check out our
work at Urban!*

- **Synthetic Data:**
Generate several synthetic datasets
- **Infrastructure:**
Creating the infrastructure (front-end, back-end, and middleware) of the validation server
- **Defining Privacy:**
Testing and assessing various privacy-preserving methodologies

**Validation
Server**

