

# Use Case Analysis of the Data Protection Toolkit

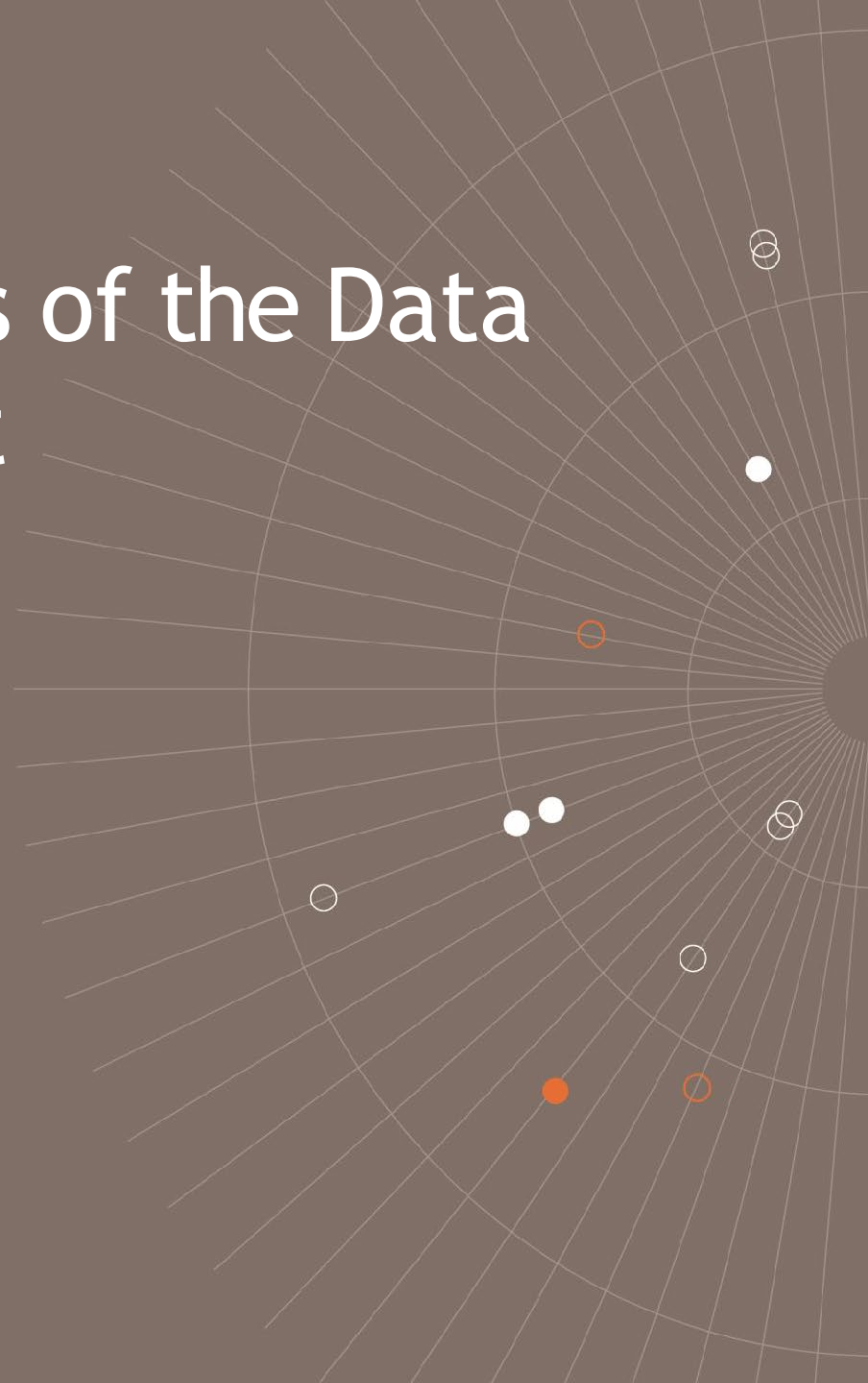
---

**October 22-24, 2024**

Michael Hawes, U.S. Census Bureau

Martha Stapleton, NORC

Team Members: Ben Reist, Seth Brohinsky,  
Peter Meyer, Julie Kubelka



# Agenda

01 Introduction

---

02 Overview of Interviews

---

03 Findings

---

04 Next Steps

---



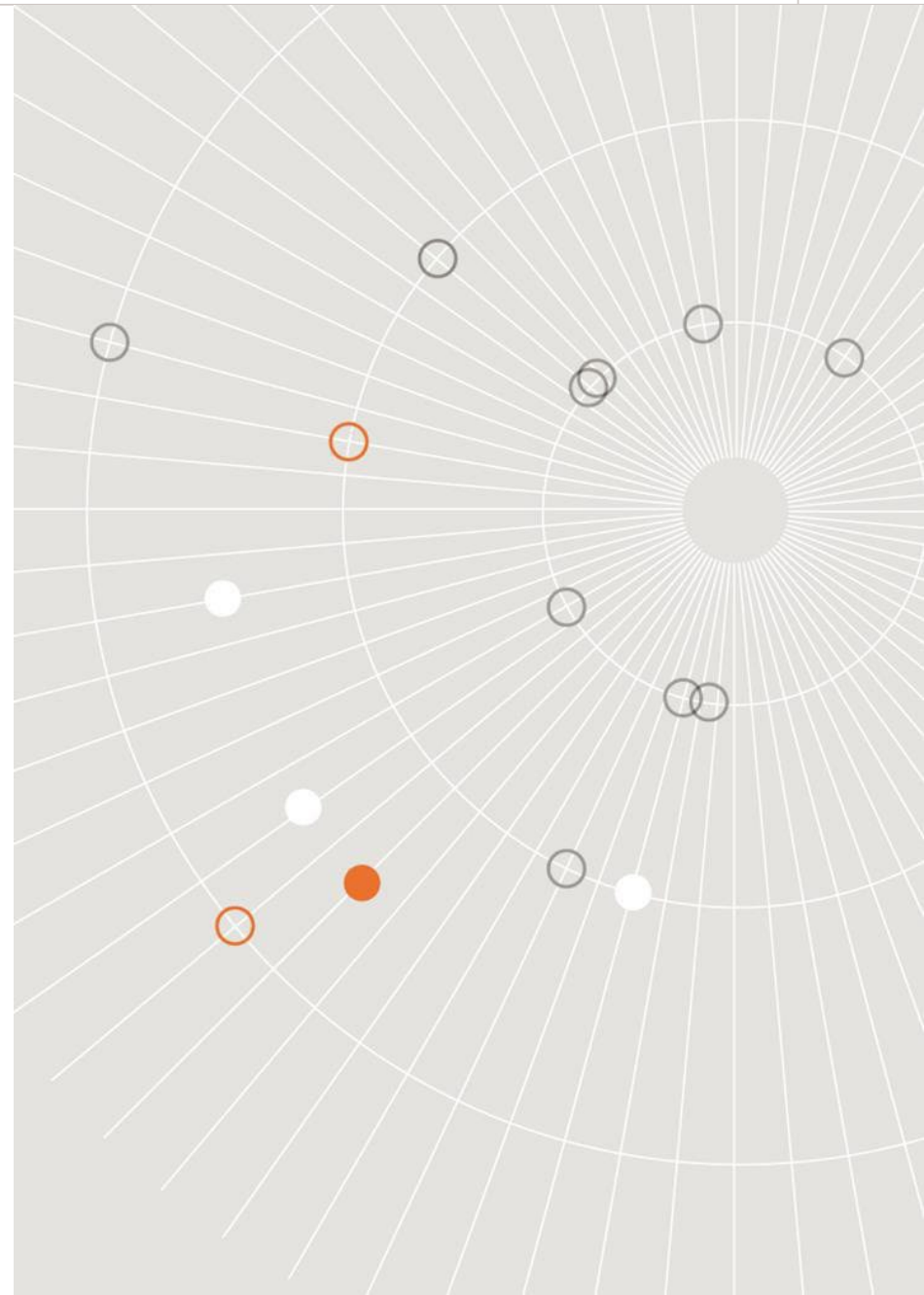
---

# Introduction: Understanding the Context



## Study Purpose

- Conduct a Use Case Analysis of the Data Protection Toolkit
- Identify successful uses and potential enhancements to the Toolkit
- Identify current data protection practices reported by federal and non-federal staff
- Receive feedback and suggested improvements to the Toolkit



# About the Data Protection Toolkit (DPT)

## Background

- Developed by the Federal Committee on Statistical Methodology (FCSM) and currently hosted by NCES
- Strong standards for data confidentiality and privacy
- Supports compliance with federal law
- Provides guidance for upholding trust in federal statistical agencies' ability to protect individual's privacy
- Includes—
  - Inventory of methods and models
  - Best Practices
  - Case Studies
  - Tools, Templates, and Automated Resources

NORC conducted 15 qualitative interviews to obtain the data reported in the final report.

### **Interview Structure**

- Conducted between September 22 and November 9, 2023
- 60-minute virtual interviews via Zoom
- NORC moderated the interviews

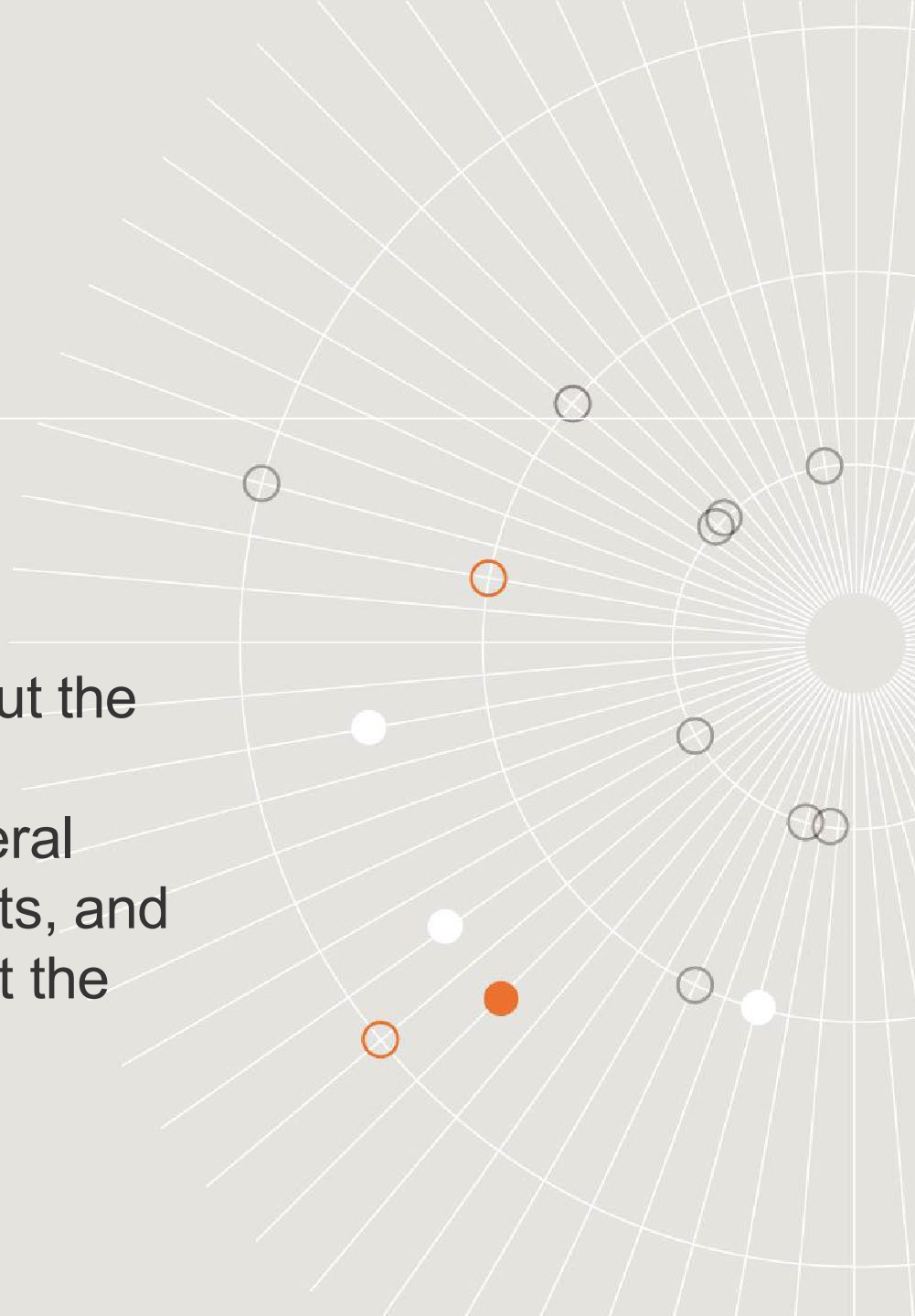
### **Participant Sample**

- NORC and NCSES collaborated to create the list of potential respondents to ensure representation across government, academia, and the private sector
- 9 respondents worked for federal agencies and 6 worked for non-federal organizations (state governments, universities, and research organizations)

---

# Findings

We begin with findings about the current data protection environment from both federal and non-federal respondents, and then move to findings about the toolkit



## The Current Data Protection Environment - Federal

Federal agencies employ multiple techniques to avoid disclosure of private and confidential information when disclosing or releasing data.

Federal agencies interviewed use a tiered access model for handling external researcher data access requests.

Two respondents noted that their data were never linked due to the siloed nature of their agency's program offices.

All federal agencies interviewed have a disclosure review process for publicly released data.

Some agencies interviewed have a formal Disclosure Review Board (DRB) to support the review process.

The federal agencies interviewed are balancing the desire to make more data available to the public, while protecting confidential information.



## The Current Data Protection Environment - Non-Federal

- Non-federal organizations employ similar techniques to avoid disclosure of private and confidential information.
- They follow formal disclosure risk review guidelines and described the resources used to guide their disclosure review process.
- They described their processes to control access to sensitive data and ensure that output is void of confidential information.
- Respondents from state governments noted the importance of meeting both state legislative requirements around data release and data confidentiality and federal collection requirements for the state-based federal surveys.
- New differential privacy practices were concerning to non-federal respondents as they were perceived as unnecessary and would ultimately reduce data usability.

## Feedback on the Toolkit - Generally Positive

Federal agency staff described the toolkit as helpful, saying it provided a source for standard information to share across agencies.

Respondents envisioned using the toolkit to learn and understand new data concepts such as differential privacy.

One respondent was using resources in the toolkit to support their work in establishing a DRB at their agency.

## Feedback on the DPT - Opportunities for Improvement



Some federal agency respondents were less sure if they would use the Toolkit as a reference because they knew the field and would turn to their colleagues with any questions.



Respondents wondered if there would be additional information in the Toolkit on different data types.



Non-federal staff noted that some of their organizations had DRBs but that the information on DRBs in the Toolkit was specific to federal DRBs.



Respondents thought that the resource list in the Toolkit was extensive and difficult to navigate.



Respondents noted that although there were different sections of the Toolkit, the intended audience and knowledge level were not explicit.

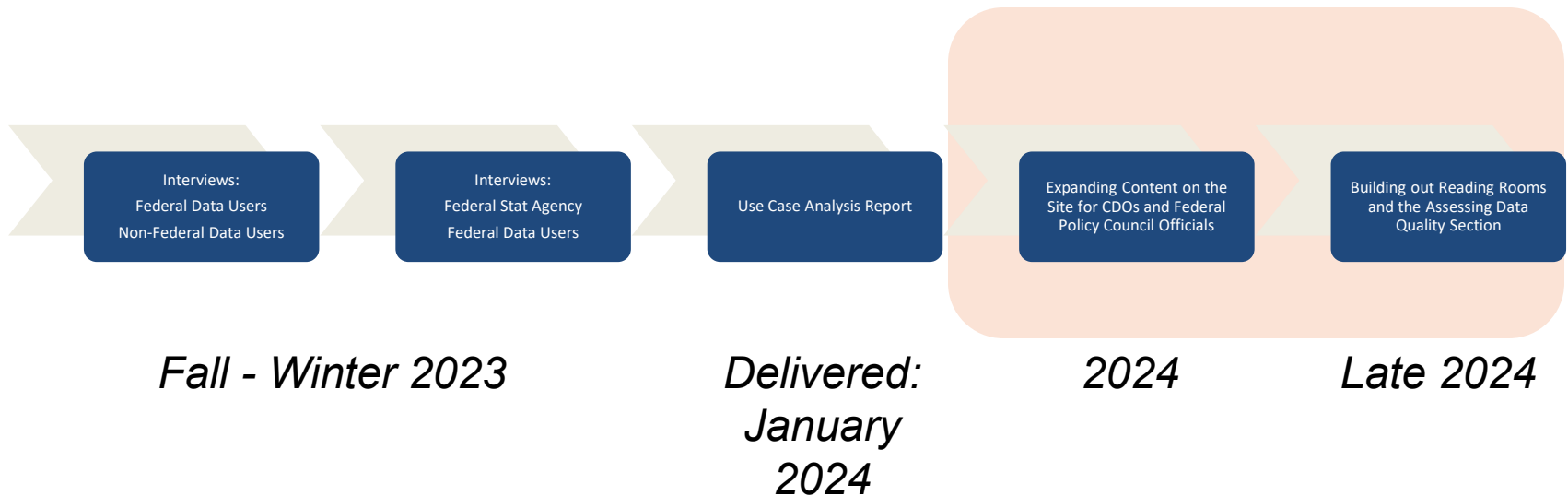


Non-federal respondents suggested specific methods to introduce the Toolkit to new audiences.



**Respondents found the DPT to be a great resource, but provided ideas to make the Toolkit more accessible and user-friendly**

# Timeline of DPT Project and Follow-up Activities



# The Future of Interagency Coordination on Data Protection

Michael Hawes  
Federal Committee on Statistical Methodology &  
Disclosure Review Officer Council

Senior Statistician for Scientific Communication  
U.S. Census Bureau

FCSM Policy and Research Conference  
October 2024

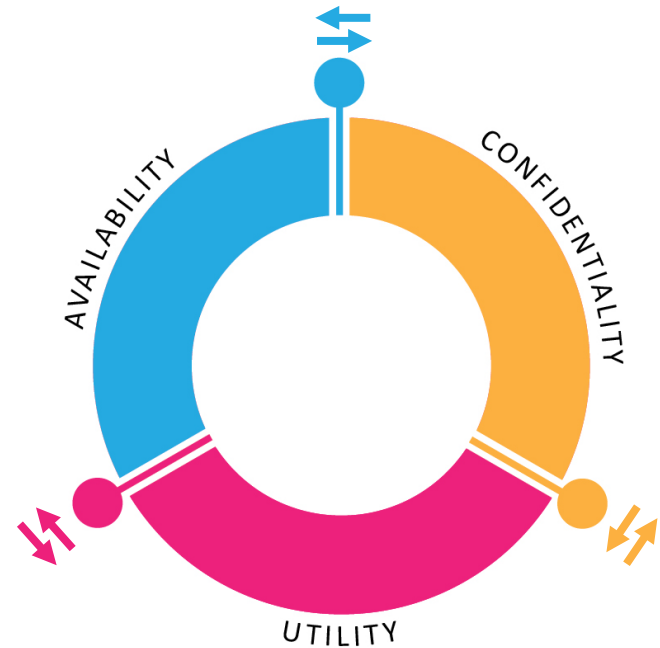
The viewpoints expressed in this presentation are the author's own, and  
not the viewpoints of the U.S. Census Bureau or the federal government.

# The Triple Trade-Off

The more statistics you publish (and the greater the granularity and accuracy of those statistics), the greater the disclosure risk.

All statistical techniques to protect confidentiality impose a tradeoff between the **degree of data protection** and the resulting **availability** and/or **utility** of the data.

You can maximize on any two dimensions, but only at profound cost to the third.



# Ensuring Fitness-for-Use

All disclosure avoidance methods, and the parameters of their implementation, impact the resulting data's fitness-for-use in different ways.

Agencies must be deliberate in their selection and implementation of disclosure avoidance methods to ensure they meet the needs of their intended data users.

Requires:

- Subject Matter Expertise
- Research and Evaluation
- Stakeholder Communication and Engagement





# Can NSDS Help?

---



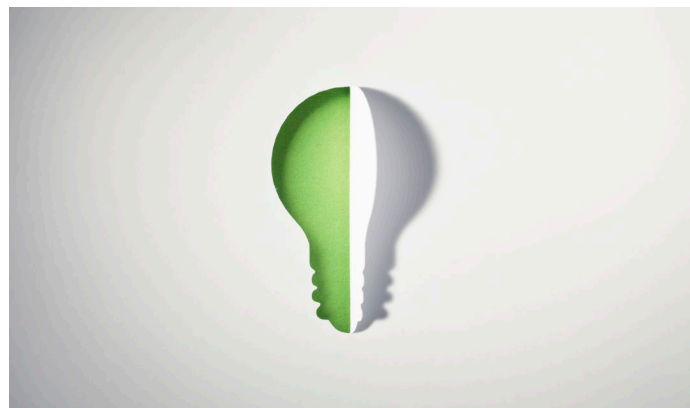


Is NSDS...

...a PLACE?



...or a PHILOSOPHY?



Agencies have largely been tackling data protection issues on their own and face numerous human capital and technological challenges.

Better **interagency coordination** could:

- Increase the pool of subject matter experts across government
- Leverage existing agency investments in data protection and tiered access solutions to benefit the entire system
- Coordinate future research and development to have broader applicability





# ANNOUNCING: The Disclosure Review Officer Council

Official Representation by the DROs from each of the Principal Statistical Agencies and Units

Tasked with:

- Information sharing (on agency initiatives, challenges, research, and collaboration opportunities)
- Crowdsourcing solutions
- Maintaining and enhancing the Data Protection Toolkit
- Developing Evidence Act implementation guidance and resources
- Establishing a government-wide community of practice

Questions?

