

Recent Advances in Data Privacy and Public Policy

Jingchen (Monika) Hu

Vassar College

FCSM 2023



In collaboration with Claire Bowen from Urban Institute.

Background

- Statistical data privacy research and applications are growing
 - Statistical agencies, from federal to city-level
 - Private companies, from Meta to Google
 - More funding from various sources specific for privacy research and applications

- Venues for presenting work and sharing ideas and community building
 - Organized sessions at FCSM, Joint Statistical Meetings (and other American Statistical Association meetings), American Association for Public Opinion Research, etc.
 - Privacy in Statistical Databases conference series
 - Many established venues are organized by and for the Computer Science community
 - A few hosted by the National Bureau of Economic Research recently

Presentation outline

- A need for venues for privacy and public policy!
- Main goal:

To improve and strengthen the relationships among the privacy experts, researchers, data stewards, data practitioners, and public policymakers

- A series of workshops between October 2022 and January 2023
 - Collaboration between the National Institute of Statistical Sciences, Penn State University, Vassar College, and Urban Institute
 - Supported by the Alfred P. Sloan Foundation
- Ongoing efforts to establish a more permanent conference on data privacy and public policy - **we welcome your involvement!**

Workshop series #1

“Census, Differential Privacy, and Civil Rights”

- One-day hybrid workshop at Penn State Institute for Computational and Data Sciences in October 2022
- 15 participants
 - Representing a cross section of academia, government, non-profit research, and civil rights groups
 - Academic backgrounds include computer science, statistics, demography, political science, law, and public policy
- Topics presented include
 - Demographic equity issues in federal data
 - Analysis of end-to-end data quality issues for Native American Populations
 - Social construction of data quality and uncertainty in official statistics

Workshop series #2

“Current Challenges and Potential Future Work”

- 1.5-hour webinar in November 2022
- Precursor to the workshop series #3
NISS IOF forum on “Advancing Demographic Equity with Privacy Preserving Methodologies”
(January 2023)
- Four panelists from
 - Academia
 - Government
 - NGO/Think Tank
 - Data users/practitioners
- 117 attendees with a full presentation of academia, government, NGO/Think Tank, and data users/practitioners



JERRY REITER
Duke University

JENNY THOMPSON
U.S. Census Bureau

TOM KRENZKE
Westat, Inc.

AMY O'HARA
Georgetown University

Workshop series #2

“Current Challenges and Potential Future Work”

- Topics discussed include each panelist
 - How they got involved in data privacy and confidentiality work
 - Describing a recent data privacy work project and its challenges and potential impact
 - Giving examples from their experience where differential privacy has been used, either successfully or with some problems
 - Evaluation of the tradeoff between privacy and utility
 - Advice for someone early or mid-career who are thinking about getting involved in this field

To view the recording: <https://www.youtube.com/watch?v=Ti-cm3XO5-A&t=2s> or search “IOF Webinar: Advancing Demographic Equity with Privacy Preserving Methodologies” on YouTube

Workshop series #3

“Advancing Demographic Equity with Privacy Preserving Methodologies”

- Two-day hybrid workshop at Urban Institute in January 2023
- Workshop goal:

To bridge the communication gap and identify what tools and educational materials are needed to improve expanding access to confidential demographic data for better public policy decision-making

- 60 in-person and 15 virtual attendees



Workshop series #3

“Advancing Demographic Equity with Privacy Preserving Methodologies”

- Six planned sessions
 - Introduction to challenges in statistical privacy
 - Applying statistical privacy methods and techniques
 - Public policy and privacy law
 - Panel on funding opportunities (representatives from Sloan Foundation, NSF, and Meta)
- Each non-panel session had 2-3 speakers with a full representation of academia, government, NGO/Think Tank, and data practitioners

Workshop series #3

“Advancing Demographic Equity with Privacy Preserving Methodologies”

- Session topics

- Privacy and utility tradeoff
- Data privacy and data equity
- Differential privacy for surveys
- Bayesian methods for differentially private synthetic data
- Use cases of creating synthetic human services datasets
- Challenges related to collection of data on antimicrobial resistance in animals
- An overview of the differential privacy framework for the 2020 Census
- Measurement error in practice for official statistics
- Case studies from the NCHS Data Linkage Program
- Data equity and confidentiality treatments
- Presentation and Q&A of the [ACDEB \(Advisory Committee on Data for Evidence Building\) Year 2 Recommendations](#)

Workshop series #3

“Advancing Demographic Equity with Privacy Preserving Methodologies”

- Working roundtables and **working groups**
 - Research challenges in statistical data privacy techniques
 - **Application challenges in statistical data privacy techniques**
 - Education and communication
 - **Equity in the context of privacy**
 - Public policy and privacy law
- Key takeaways
 - The need for a centralized resource of data privacy methods and best practices
 - More applied and practical applications of data privacy methodologies
 - **Events that bring together people from diverse backgrounds, such as workshops and conferences**

Ongoing efforts

- To establish a more permanent conference series under the theme of **Privacy and Public Policy**, a first of its kind
- 5 co-chairs (academia, government, and NGO/Think Tank) on the program committee and planning committee
 - Valbona Bejleri (USDA NASS; valbona.bejleri@usda.gov)
 - Claire Bowen (Urban Institute; CBowen@urban.org)
 - V. Joseph Hotz (Duke University; v.joseph.hotz@duke.edu)
 - Quentin Brummet (NORC; Brummet-Quentin@norc.org)
 - Monika Hu (Vassar College; jihu@vassar.edu)



Ongoing efforts

- Current conference details
 - September 13-14, 2024
 - Massive Data Institute at Georgetown University, Washington D.C.
 - Single-track conference, i.e., no concurrent sessions
 - Invited panel under the tentative theme of Translating Privacy Law to Practice
 - Poster session
 - Abstract submission (2 pager) with reviewer process; deadline February 29, 2024

- We welcome your involvement!
 - Mark your calendar
 - Submit an abstract
 - Share amongst your colleagues

Thank you! Questions?

Jingchen (Monika) Hu

jihu@vassar.edu

Vassar College