Navigating the Data Use & Privacy Tradeoff:
An Information-Theoretic Approach to Privacy Regulation

Ayelet Gordon-Tapiero (Law Faculty & Dept. of Computer Science, Hebrew University);
Katrina Ligett (Dept. of Computer Science, Hebrew University);
Kobbi Nissim (Dept. of Computer Science, Georgetown University);
Muhammad Saad (McCourt School, Georgetown University)

Privacy & Public Policy Conference, Georgetown University
February 9-10, 2026

Research Motivation

★ **Practical Motivation:**
Data privacy regulation should account for *disclosure risks* that *accumulate* from *multiple uses of data*, and *across parties*, *sources* and *timelines.*

★ **Theoretical Motivation:**
Understand *economic* and *information theoretic* properties of data use and privacy

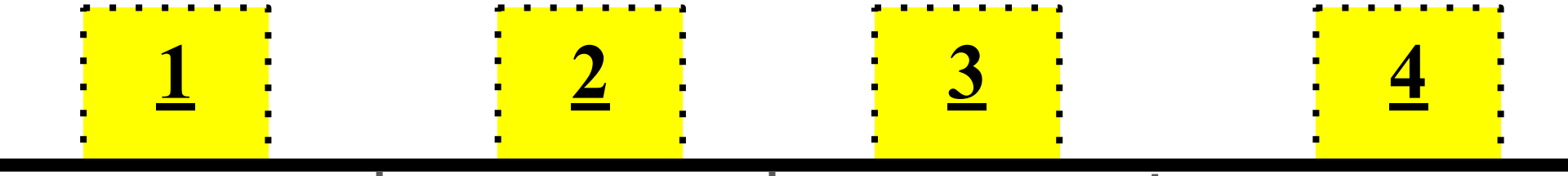
Current Regulation

★ U.S. regulation recognizes that privacy harms accumulate with use of data

★ **Mosaic Effect**: Individual data releases may be harmless, but could lead to privacy disclosure when viewed with other sources

★ **Unit of regulation**: Individual dataset at the point of release

Contributions



Develop general theory of data use as a central concept and how privacy harms accumulate

Apply information theory, economics, and privacy-related mathematical frameworks

Demonstrate even if two data uses individual preserve privacy, they could together lead to a breach

Show that current legal approaches to privacy do not fully recognize accumulation of privacy harms

Regulatory Perspective:

- Coordinate data use across users, sources and timeline instead of individual releases
- Use programmable mathematical frameworks about privacy to coordinate how data is used and privacy depleted
- Regulate data use as a rival good. Privacy depletes with multiple data uses, hence, need to manage the scarcity.

Regulatory Principles:

Track/coordinate across different uses/users:
Manage cumulative privacy loss through a shared and targeted privacy budget

Reduction of Data Use:
Prioritize low privacy-cost uses, regulate access and accept lower accuracy in outputs

Use Synthetic Data:
Synthetic data generated with mathematical guarantees can be reused without privacy loss

Institutional Coordination:
Appoint data curators to track, allocate, and enforce privacy budgets across users