SaTC: CORE: Large: Collaborative: **Accountable Information Use: Privacy and Fairness in Decision-Making Systems: SoK: Differential Privacy as a Causal Property**

Michael Carl Tschantz (ICSI) and Anupam Datta (CMU) with Shayak Sen (AI Lens)

https://fairlyaccountable.org/satc/

Differential Privacy is better understood as a bound on a causal effect size than as a bound on associations, correlations, or knowledge.

Associational View

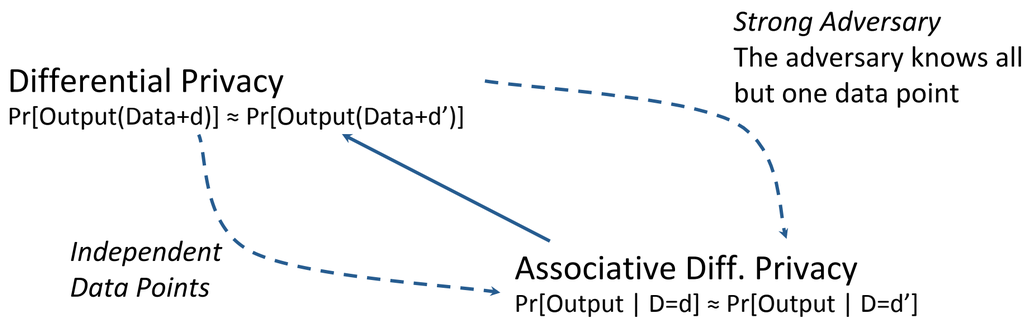
Bayesian DP, Dependent DP,
statistical nondisclosure

1. Pr[Algorithm(data+d)=o] ≈ Pr[Algorithm(data+d')=o]
2. Pr[O=o | Database=data+d] ≈ Pr[O=o | Database=data+d']
   where O = Algorithm(Database)
3. Pr[O=o | D=d] ≈ Pr[O=o | D=d']
   where O = Algorithm(Database) and Database = data+D

How much does the adversary know about a data point from seeing the output?

Hard (almost impossible in general)

"Requires implicit assumptions":



Disallows "Smoking causes cancer"

Prevents most things, good and bad

Causal View

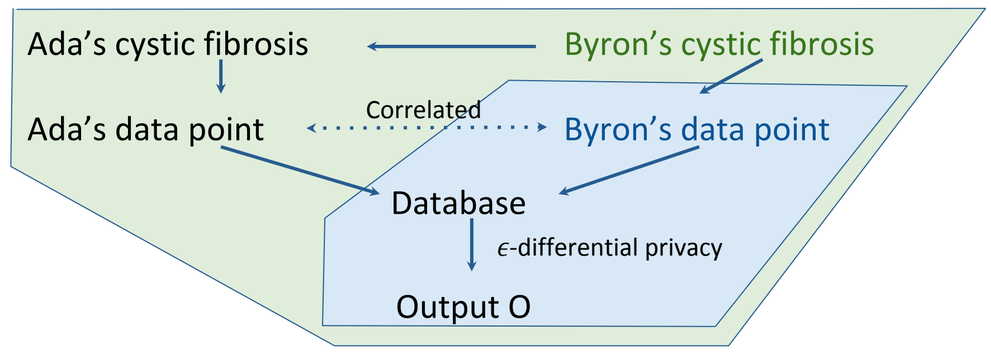
Differential privacy, semantic privacy, noninterference, semantic security

Pr[O=o | ***do*** D=d] ≈ Pr[O=o | ***do*** D=d']
where O = Algorithm(Database) and
Database = data+D

How much does what the adversary learn change with a data point?

Add noise

About data points, not underlying attributes:



Allows Gaydar-like studies

Enough to encourage participation: going to release results either way

Key Outcomes

Explains common misconconception:
- association doesn't imply causation
- no causation doesn't imply no association
- DP doesn't limit association (information)

Explains why DP is used outside of privacy

The grant trained a grad student/postdoc

To appear at IEEE S&P 2020

Ongoing Work

Privacy and nondiscrimination are related

	Privacy	Nondiscrimination
Causal	Differential privacy	Something like Disparate treatment
Associative	Statistical nondisclosure	Something like Disparate impact