Computing over Distributed Sensitive Data

Stephen Chong, Marco Gaboardi, James Honaker, Kobbi Nissim, Salil Vadhan

Challenge:

Different organizations collect similar data or data about similar populations. When the data is sensitive, can we benefits from analyzing this data while avoiding the need to share them between the different entities?
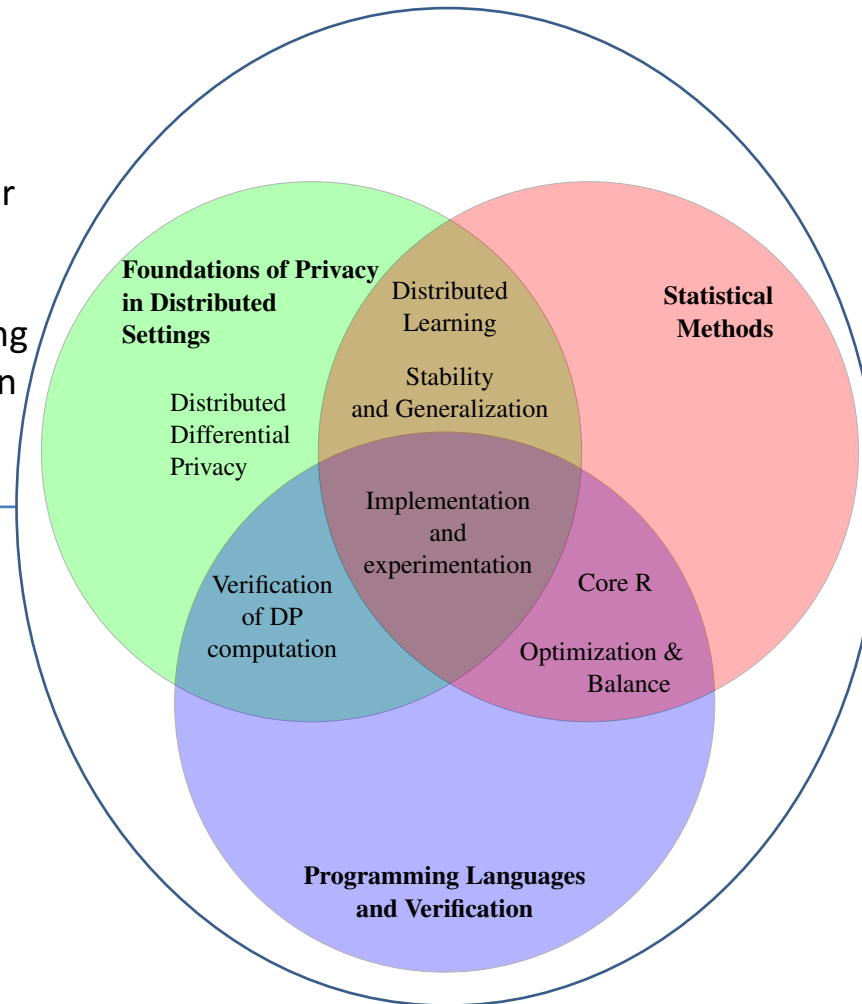
Solution:

- Distributing differential privacy computations.
- Sharing the computation (and not the data) across multiple entities.

Foundations of Privacy in Distributed Settings

Distributed Differential Privacy

Distributed Learning

Stability and Generalization

Statistical Methods

Implementation and experimentation

Verification of DP computation

Core R

Optimization & Balance

Programming Languages and Verification

Scientific Impact:

A conceptual-foundational understanding of distributed computations over sensitive data

A collection of tools that will achieve many benefits of data sharing without data owners having to share the data.

Broader Impact:

- Understanding the tradeoffs between privacy and utility in distributed data,
- Helping policy makers to take informed decisions in the already ongoing negotiation of privacy,
- development of new curricula material and training