

Project Title: Statistics-Friendly Formal Privacy for Establishment Data

Awardee: Penn State University

Contract Number: 49100421C0022

Award Type: R&D

Award Amount: \$484,525

Principal Investigator: Daniel Kifer

NSF Funding Program: National Center for Science and Engineering Statistics

NSF Program Director: Jennifer Sinibaldi, jsinibal@nsf.gov

Abstract:

The goal of this project is to develop a customizable suite of privacy technology, inspired by differential privacy, that is suitable for statistical use of establishment data. We plan to (1) create formal privacy definitions that can be tailored to the unique properties of business data and the unique needs of different statistical agencies, (2) develop provably correct algorithms for protecting establishment data under the customized privacy definitions, and (3) develop statistical methodology for analysis of the released statistics. The privacy definitions must meet modern standards of transparency, composition, postprocessing immunity and privacy loss accounting. They must also be "statistics-friendly" to ensure that privacy-protected data is usable for statistical purposes (e.g., uncertainty estimation). Our plan is to target inter-related statistics programs, with the Current Employment Statistics (CES) and Quarterly Census of Employment and Wages (QCEW) being our motivating applications. Since QCEW is used to benchmark estimates from the CES, this provides a natural ecosystem for evaluating the fitness-for-use of privacy-preserving establishment statistics.