# Re[DP]: Realistic Data Mining Under Differential Privacy

Challenge:

- Differential privacy: a gold standard for analysis of sensitive data
- Adoption by industry & government remains difficult.
- How to close gap between theory and practice?

Solution:

- Focus on under-studied aspects of data mining workflows
- Simplify algorithm design and deployment
- Design robust benchmarks for principled evaluation

World according to DP research

Real World

DPComp.org
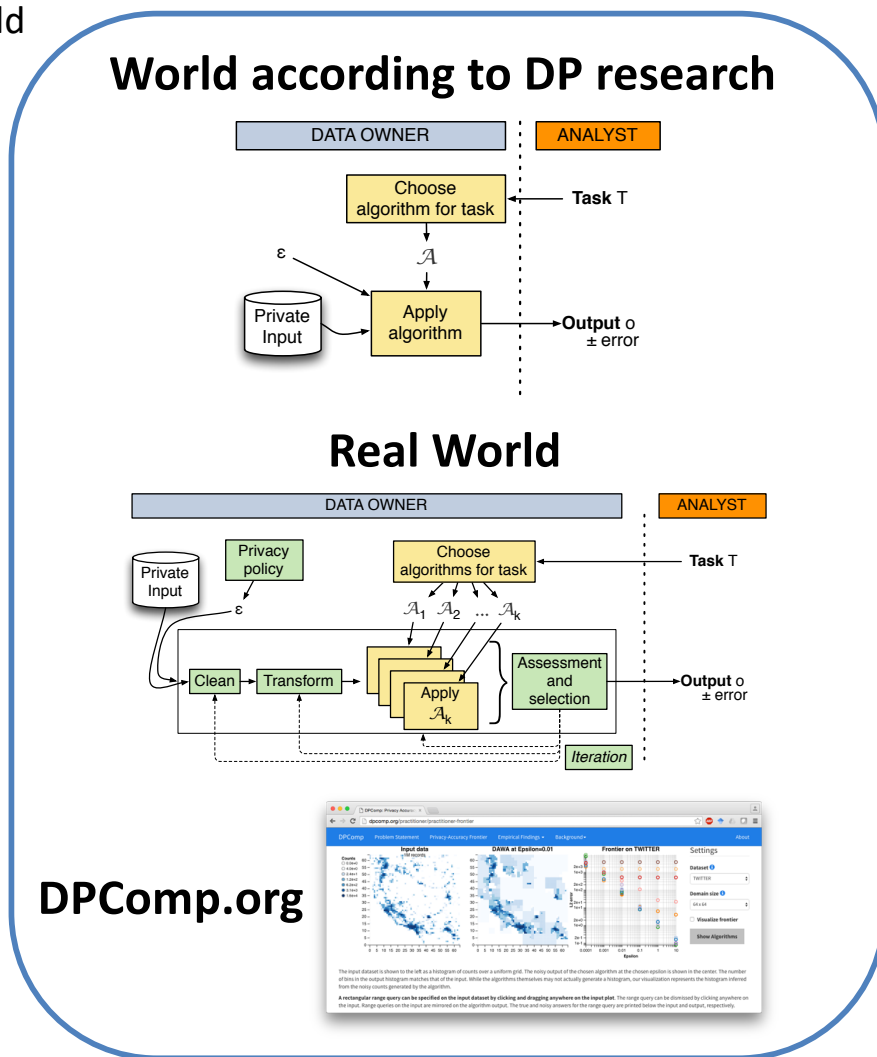
Scientific Impact:

- End-to-end privacy for realistic data mining workflows
- Sound and reproducible methodology for evaluation
- Novel error-optimal algorithms for data mining tasks, including PrivateSQL for answering complex queries privately.
- The Ektelo platform for designing and deploying privacy algorithms.

Broader Impacts:

- Expand practical usefulness of privacy algorithms to increase data sharing.
- DPComp.org: online resource to assess state-of-the-art of differentially private algorithms
- Real-world deployment: tech-transfer to US Census Bureau
- Open source code: DPComp, Ektelo, HDMM, and NIST DP contest-winning solution.
- Education and outreach
  - Tutorials at VLDB 2016 and Simons
  - Privacy Bootcamps for Census and other government agencies.

PIs: Michael Hay (Colgate), Ashwin Machanavajjhala (Duke), Gerome Miklau (UMass Amherst)
Award #s: 1409125, 1409143, 1408982

Colgate   Duke UNIVERSITY   UMassAmherst