Usable and Secure Data Analysis

Mayank Varia, Boston University, varia@bu.edu

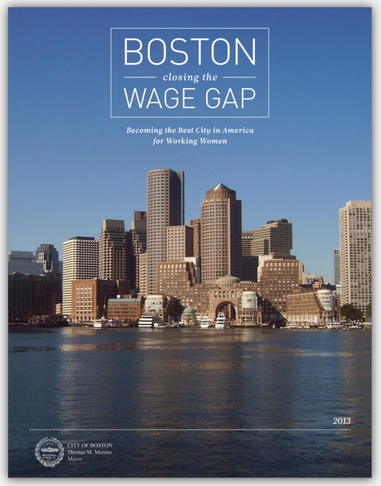
BU Initiative on Cities

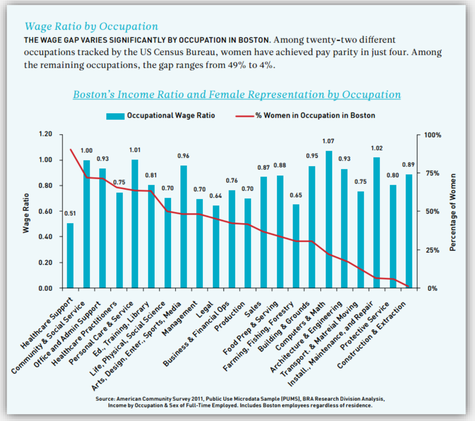
MODULAR APPROACH TO CLOUD SECURITY

Problem

aggregate sensitive salary data, calculate pay equity

Meeting with Mayor Menino at BU, 7/31/14





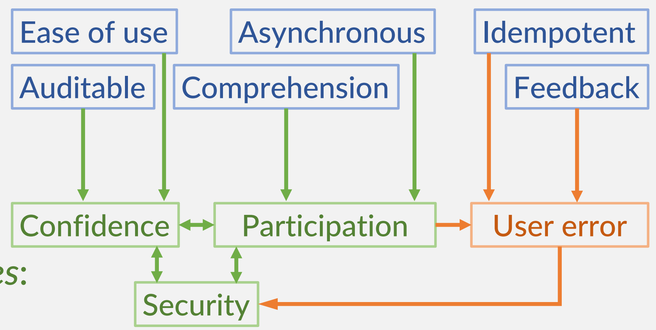
GOAL 3

Evaluating Success

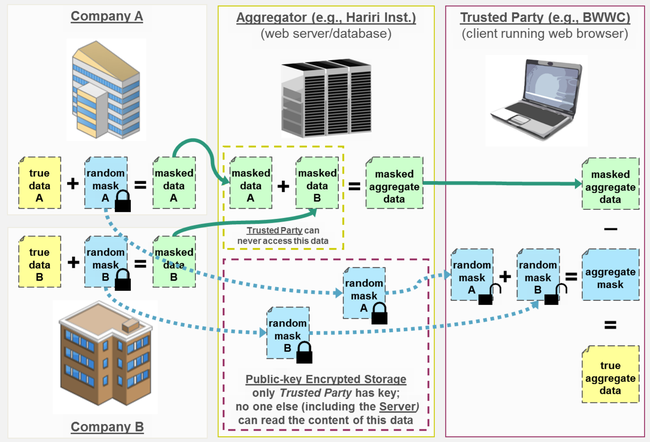
Employers agree to participate in a biennial review to discuss successes and challenges, as well as contribute data to a report compiled by a third-party on the Compact's success to date. Employer-level data would not be identified in the report. The specific data to be reported will build on data already required by federal and state authorities and should not create an additional reporting burden.

Usability

increases confidence, participation, and security



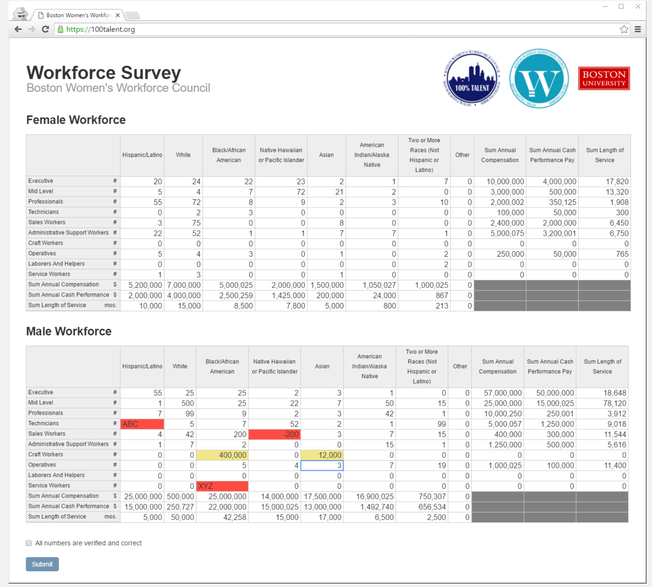
Methodology & Interface



Contributors only learn what the trusted party tells them.

The aggregator never learns anything. Neither does any hacker attacking it.

Trusted party only learns analytics that were authorized by contributors.



Reaction

The Boston Globe

The congresswoman, who had signed onto a bill addressing income disparity between men and women, was impressed by the relevance he outlined. *"It's linking it back for the members of Congress,"* Clark said. *"Nobody would think, oh, the Paycheck Fairness Act, how is that tied into NSF funding?"*

On Point with Tom Ashbrook

BWWC co-chair Evelyn Murphy on secure multi-party computation: *"It's used in computer science applications, but it has never been used for public good. Here, we're beginning to show how to use this sophisticated computer science research for public programs."*

Extensions

1. Scale to big data
2. Integrate with legacy code

Setup	Runtime
Insecure baseline	16.2 min
Our system	*17.5 min*
Secure baseline	> 120 min

Test: compute market concentration on 156 GB of NYC cab trip data with 5 Amazon Web Services machines

More info

Collaborators

- Azer Bestavros
- Frederick Jansen
- Andrei Lapets
- Malte Schwarzkopf
- Nikolaj Volgushev

References

1. User-centric Distributed Solutions for Privacy-preserving Analytics, *Comm. of the ACM, 2017*
2. Secure Multi-Party Computation for Analytics Deployed as a Lightweight Web Application, *SecDev 2016*
3. Demo: Integrating MPC in Big Data Workflows, *CCS 2016*
4. Programming Support for an Integrated Multi-Party Computation and MapReduce Infrastructure, *BU CS Technical Report*
5. Scather: Programming with Multi-party Computation & MapReduce. *BU CS Tech Report*