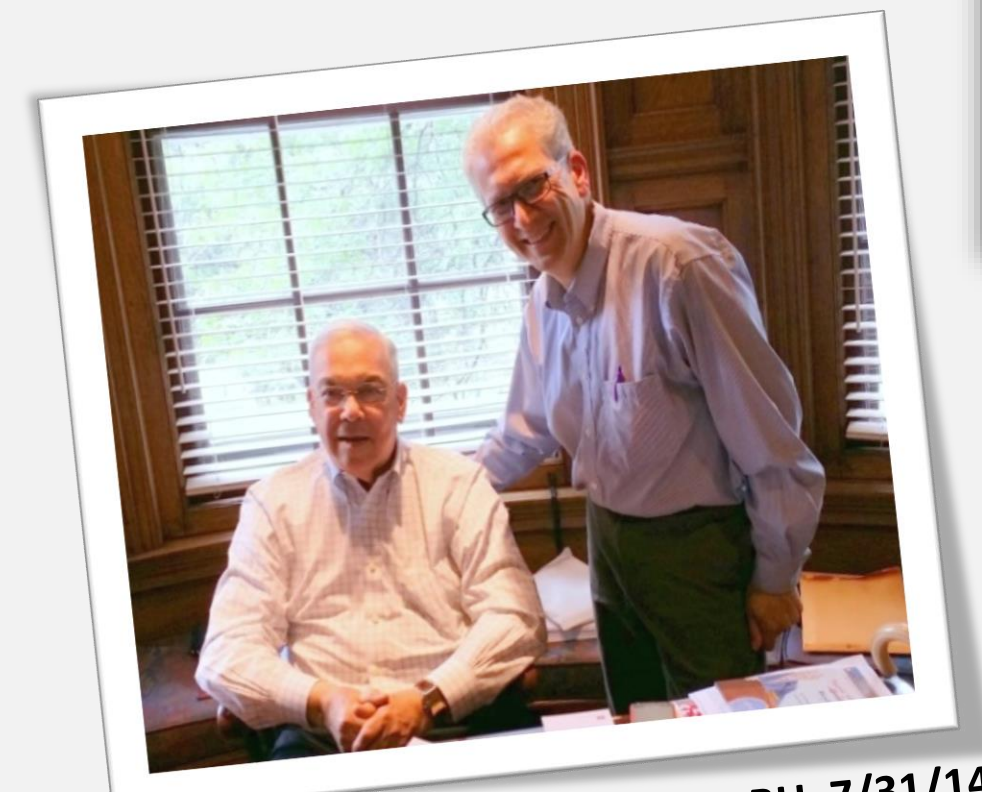
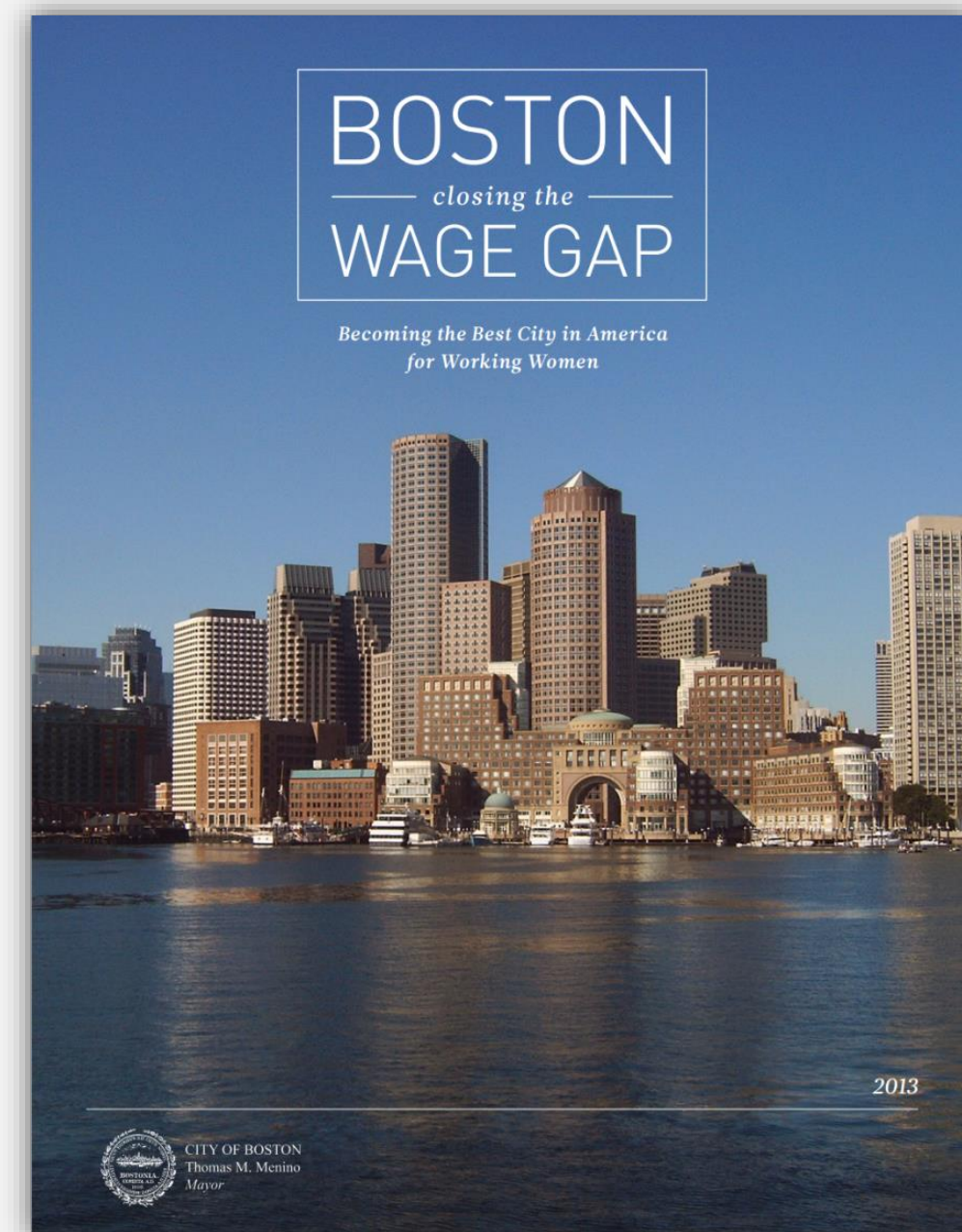


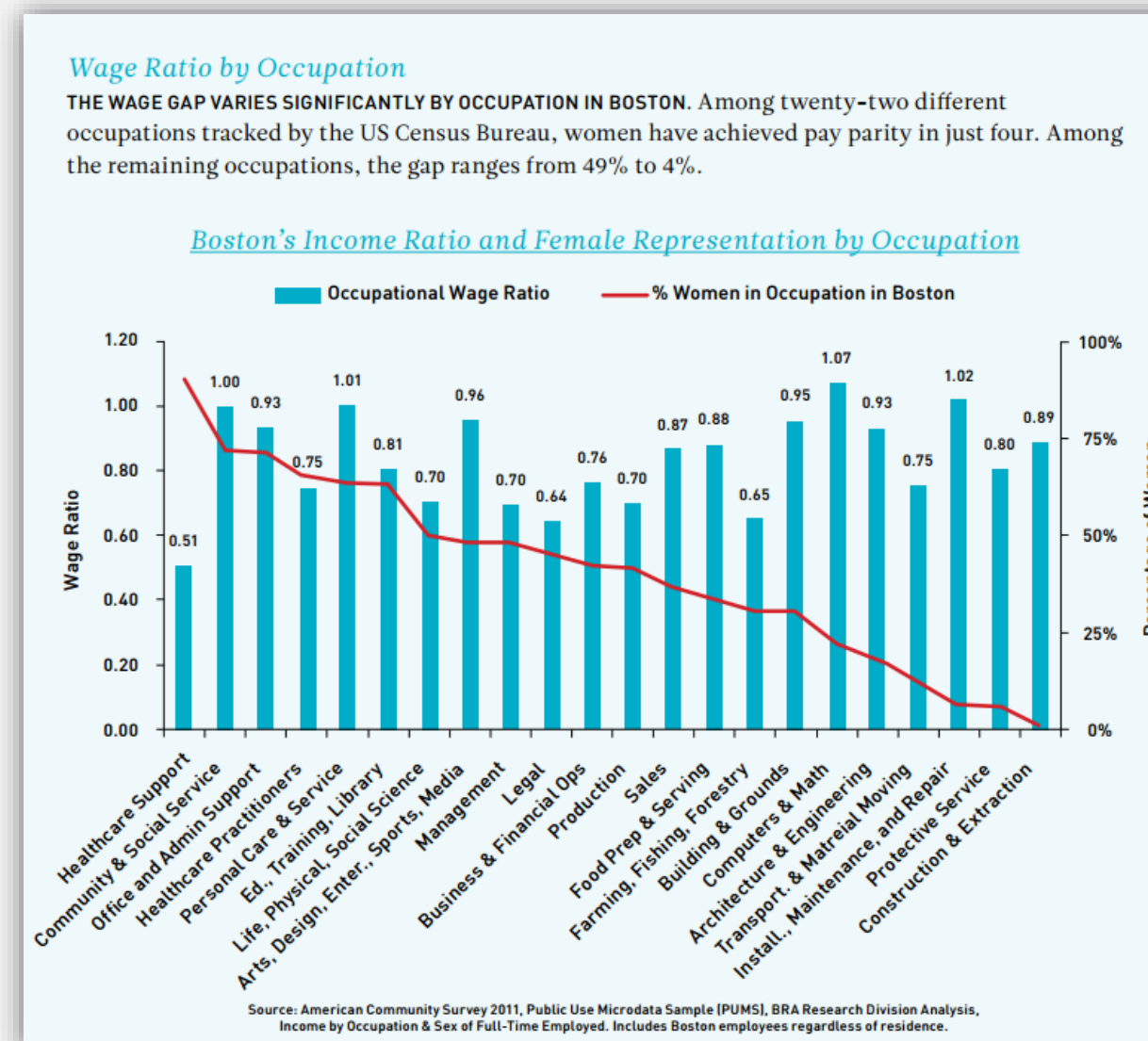
# Usable and Secure Data Analysis

Mayank Varia, Boston University, varia@bu.edu

**Problem**  
aggregate sensitive salary data, calculate pay equity



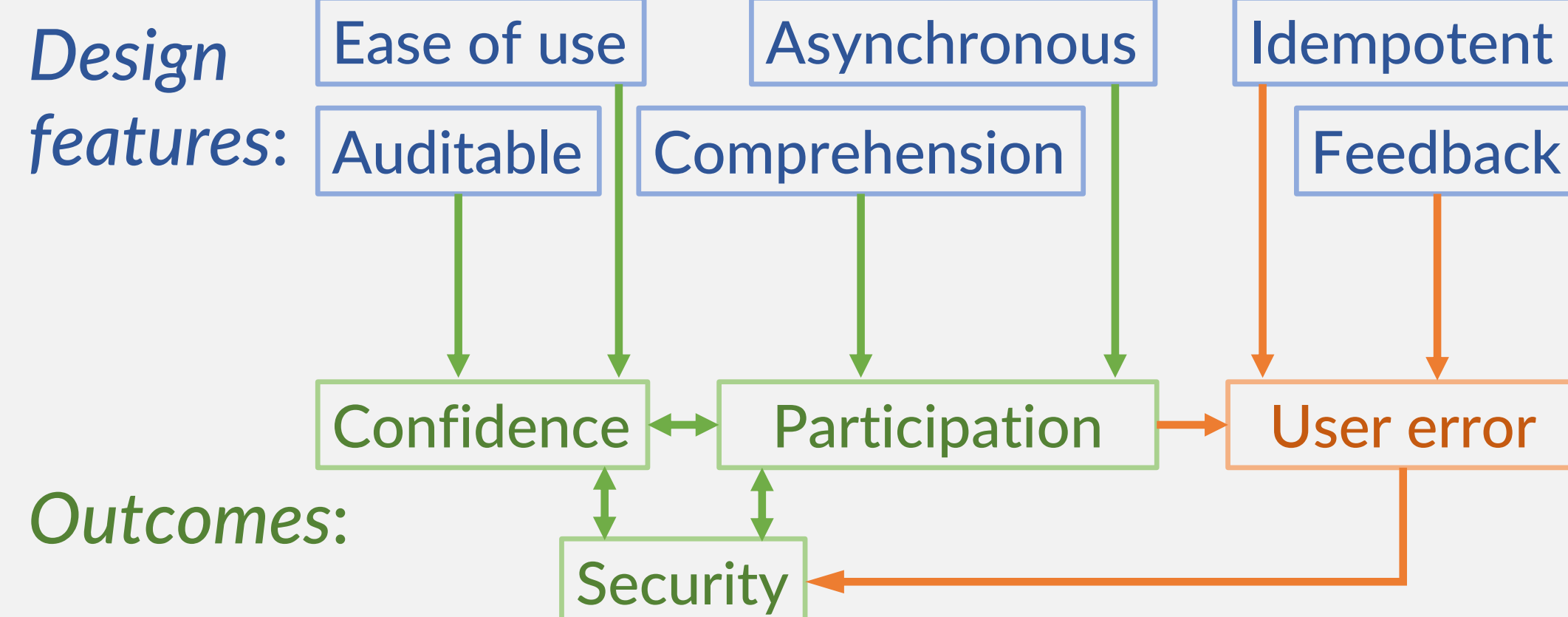
Meeting with Mayor Menino at BU, 7/31/14



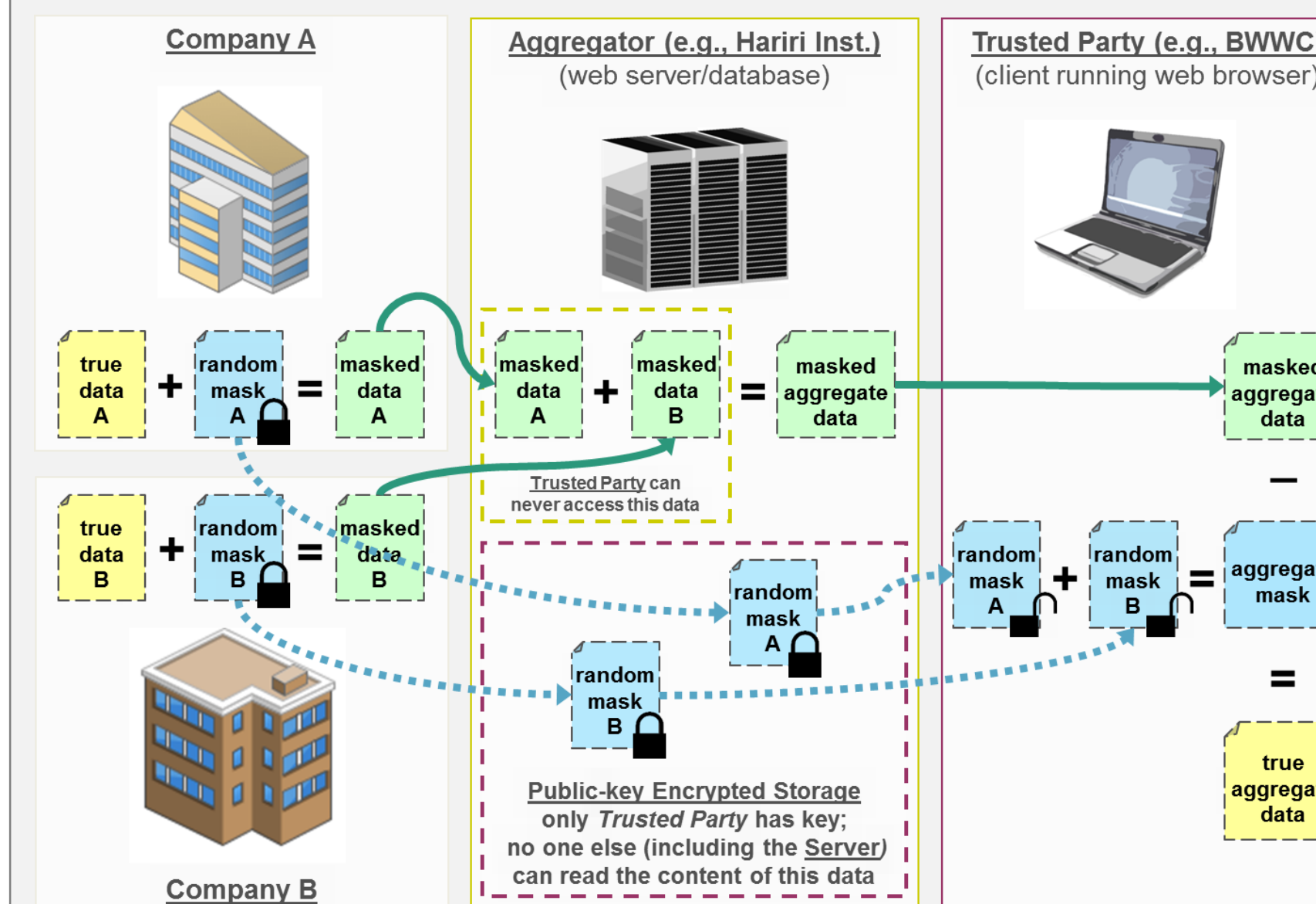
**GOAL 3**  
Evaluating Success

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.

**Workforce Survey**  
Boston Women's Workforce Council

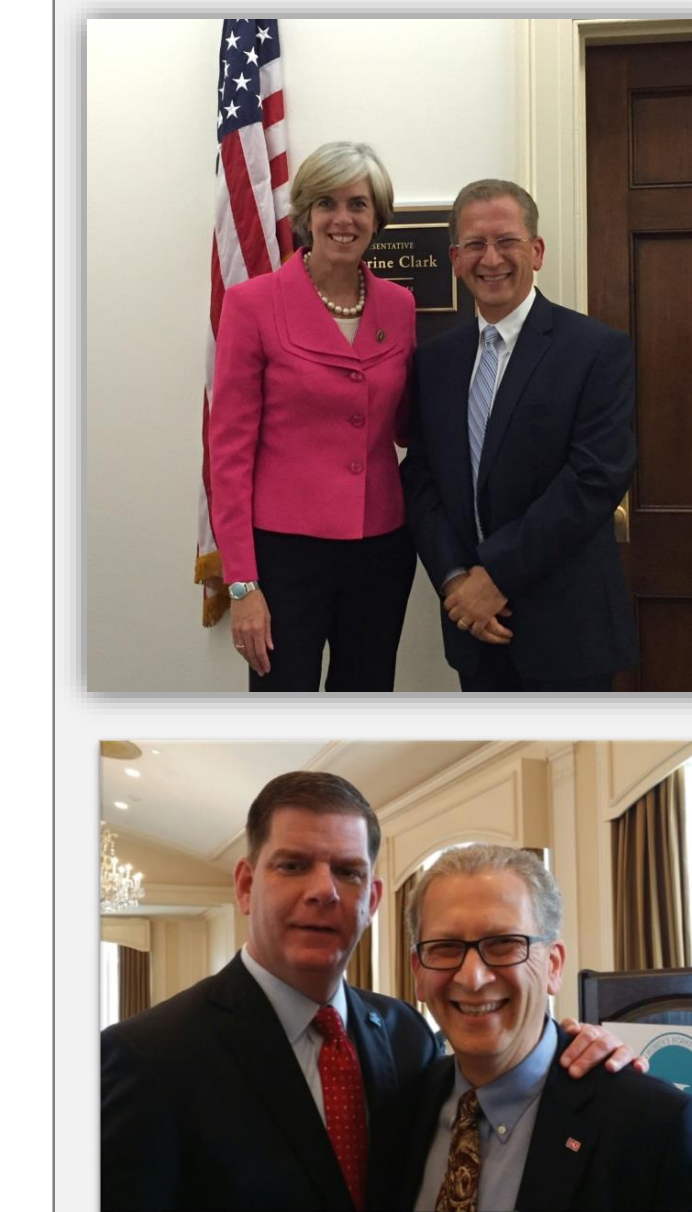
**Female Workforce**

	Hispanic/Latino	White	Black/African American	Native Hawaiian or Pacific Islander	Asian	American Indian/Alaska Native	Two or More Races (Not Hispanic or Latino)	Other	Sum Annual Compensation	Sum Annual Cash Performance Pay	Sum Length of Service
Executive	20	24	22	23	2	1	7	0	10,000,000	4,000,000	17,820
Mid Level	5	4	7	72	21	2	0	0	3,000,000	500,000	13,320
Professionals	55	72	8	9	2	3	10	0	2,000,002	350,125	1,908
Technicians	0	2	3	0	0	0	0	0	100,000	50,000	300
Sales Workers	3	75	0	0	8	0	0	0	2,400,000	2,000,000	6,450
Administrative Support Workers	22	52	1	1	7	7	1	0	5,000,075	3,200,001	6,750
Craft Workers	0	0	0	0	0	0	0	0	0	0	0
Operatives	5	4	3	0	1	0	2	0	250,000	50,000	765
Laborers And Helpers	0	0	0	0	0	0	0	0	0	0	0
Service Workers	1	3	0	0	1	0	0	0	0	0	0
Sum Annual Compensation	5,200,000	7,000,000	5,000,025	2,000,000	1,500,000	1,050,027	1,000,025	0			
Sum Annual Cash Performance	2,000,000	4,000,000	2,500,250	1,425,000	200,000	24,000	867	0			
Sum Length of Service	10,000	15,000	8,500	7,800	5,000	800	215	0			

**Male Workforce**

	Hispanic/Latino	White	Black/African American	Native Hawaiian or Pacific Islander	Asian	American Indian/Alaska Native	Two or More Races (Not Hispanic or Latino)	Other	Sum Annual Compensation	Sum Annual Cash Performance Pay	Sum Length of Service
Executive	55	25	25	2	3	1	0	0	57,000,000	50,000,000	18,648
Mid Level	1	500	25	22	7	50	15	0	25,000,000	15,000,025	78,120
Professionals	7	99	9	2	3	42	1	0	10,000,250	250,001	3,912
Technicians	5	7	52	2	1	99	0	0	5,000,057	1,250,000	9,918
Sales Workers	4	42	200	-200	3	7	15	0	400,000	300,000	11,544
Administrative Support Workers	1	7	2	0	0	15	1	0	1,250,000	500,000	5,616
Craft Workers	0	0	400,000	0	12,000	0	0	0	0	0	0
Operatives	0	0	5	4	3	7	19	0	1,000,025	100,000	11,400
Laborers And Helpers	0	0	0	0	0	0	0	0	0	0	0
Service Workers	0	0	0	0	0	0	0	0	0	0	0
Sum Annual Compensation	25,000,000	500,000	25,000,000	14,000,000	17,500,000	16,900,025	750,307	0			
Sum Annual Cash Performance	15,000,000	250,727	22,000,000	15,000,025	13,000,000	1,492,740	656,534	0			
Sum Length of Service	5,000	50,000	42,258	15,000	17,000	6,500	2,500	0			

## 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?"

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

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