Improving the Privacy and Security of Online Survey Data Collection, Storage, and Processing (CNS-0627488)

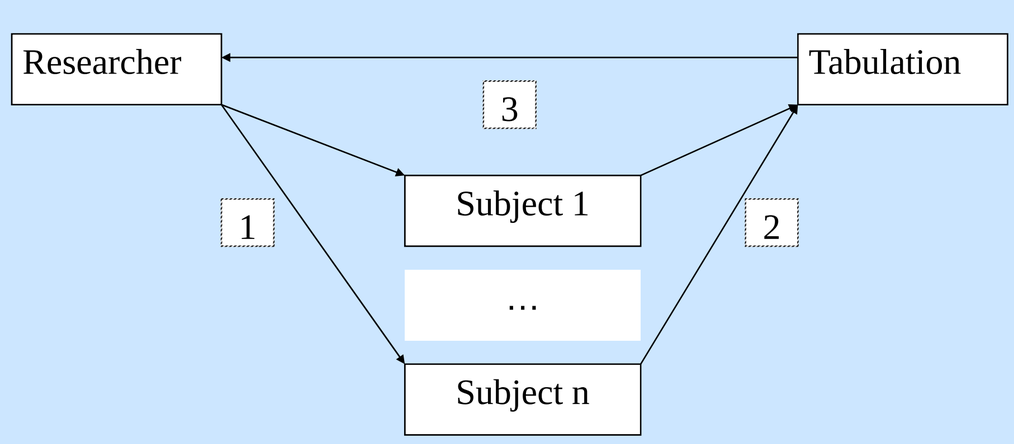
Mikhail Atallah, Juline Mills (Purdue University) www.cs.purdue.edu/homes/mja/

Objectives

Important public policy conclusions are drawn from surveys, yet problems exist in the way such data is collected and handled, resulting in low quality (biased) responses from participants. The reasons for biased responses include many (like social desirability bias, and acquiescence bias) that would be mitigated by a technology that hides the individual responses and reveals only the aggregate outcomes of the agreed-upon data analyses.

The goal of this project is to develop techniques that protect the privacy of survey participants' individual responses, and that achieve purpose-enforcement -- the requirement that the collected data can be used only for the purpose for which it was collected.

The project will make possible better privacy and security for online surveys, and result in higher quality survey responses



- Responses not viewable by researcher
- Tabulating agency learns nothing
- Researcher learns only stat aggregates
- Break-in does not compromise responses

Achieving better privacy and security, and higher-quality survey responses

Approach and Impact

New approach

- Protocols for off-line entry and analysis of survey data
- New techniques for computing with hidden data

Research Impact

- New survey technologies
- Better privacy
- Higher quality responses

- Simple statistics (mean, variance, …)
- Outlier removal
- Data sampling
- t-tests
- Chi square
- Regression (simple, multiple)
- Correlations
- ANOVA / MANOVA
- Structural equation modeling
- Factor analysis
- …done without seeing inputs

Data analysis protocols for this hidden-data framework

- Eliciting trust on the part of a survey user
  - Essential for achieving higher quality of responses
- User interface
  - Factors that help most, least (even detract)
- Conveying to non-technical survey respondent the technical high security and privacy characteristics of the system
  - Explain differences from less secure designs
- Experimental validation of the design
  - Case studies
  - Survey on Internet addiction

Human factors issues

National Science Foundation
WHERE DISCOVERIES BEGIN

NSF Cyber Trust Principal Investigators Meeting
March 16-18, 2008
New Haven, CT

PURDUE UNIVERSITY