

# CERIAS

the center for education and research in information assurance and security

## Improving the Privacy and Security of Online Survey Data Collection, Storage, and Processing

Faculty: Mikhail Atallah, Juline Mills

Students: Marina Blanton, Celestino Ruffini, Richard Wartell

### 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.*

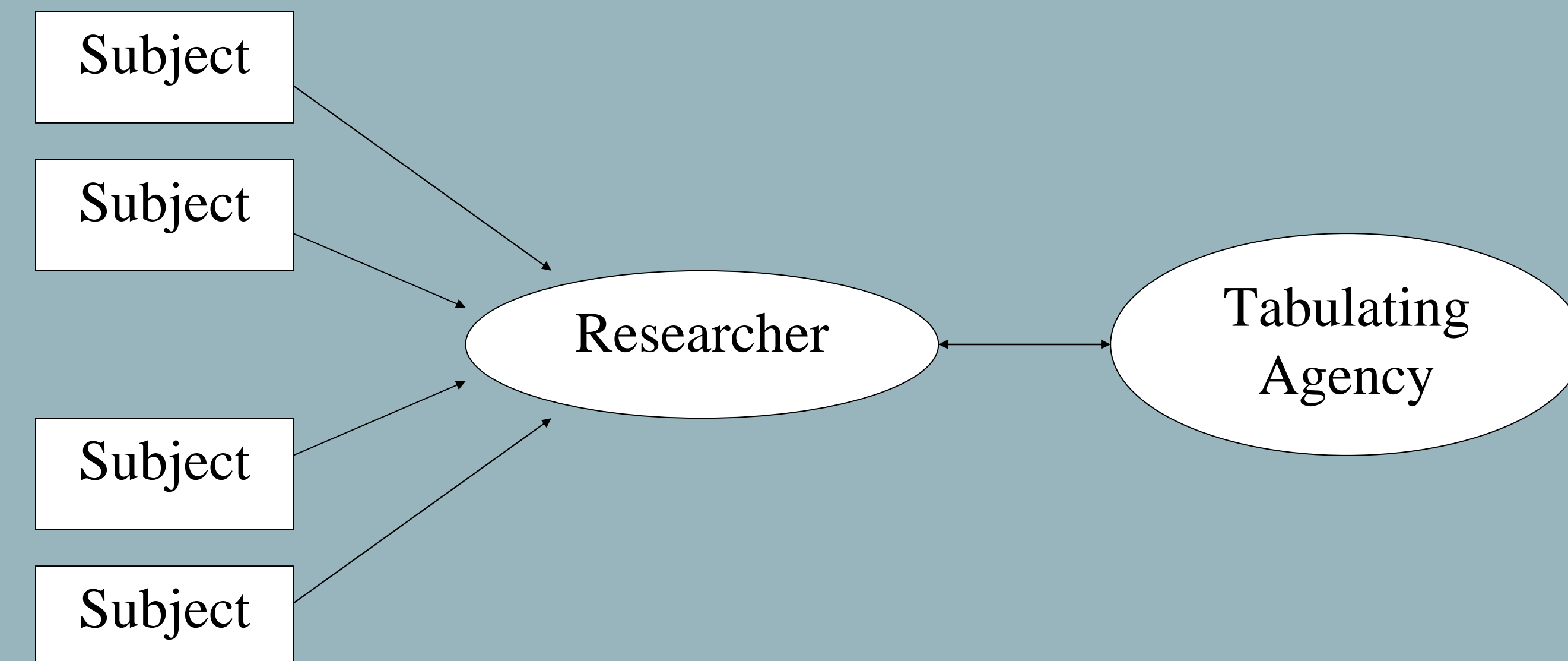
The project will make possible better privacy and security for online surveys, and result in higher quality survey responses

#### Approach and Impact

- Multi-disciplinary (CS, social sciences)
- Protocols for data entry and analysis that do not require participants to be simultaneously online
- Computing with hidden data
- Enable complex data analyses in this framework
- Change the way surveys on sensitive subjects are carried out
- User Interface for communicating privacy information

- Simple statistics (mean, variance, ...)
- Outlier removal
- Data sampling
- t-tests
- Chi square
- Regression (simple, multiple)
- Correlations
- ANOVA / MANOVA
- Structural equation modeling
- Factor analysis
- ...done by tabulating agency w/o seeing inputs

Data analysis in this hidden-data framework



- Subject responses not viewable by researcher
- Tabulating agency processes data in hidden form (does not see it, does not see outcomes of analysis)
- Researcher can view only aggregate outcomes of data analyses
- Break-in does not compromise subject's responses

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

- 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 Factor Issues