Incentives for Data Producers to Create 'Archive-Ready' Data Sets

Margaret Hedstrom, PD/PI

School of Information

University of Michigan

Research Objectives

- What incentives are most effective in convincing data producers to create data sets that are easy to archive and share?
    - What is easy/difficult effective/ineffective with current practice?
    - What types of incentives would increase compliance?
    - Which aspects of current practice should be redesigned?

Archive-ready

- Archive-ready: *data sets that meet the submission requirements of an archive under either general guidelines or a specific agreement between an archive and a data producer*
  - Data quality
  - Documentation
  - Agreements/contingencies: personal information, intellectual property, etc.

Project Participants and Partners

University of Michigan

School of Information

– Margaret Hedstrom, Assoc. Professor, PD/PI

– Yan Chen, Assoc, Assoc. Professor, co-PI

– Jingfang Niu, PhD student

ICPSR -- National Archive of Criminal Justice Data

– Myron Gutmann, Director, co-PI

– Dr. Janet Samatel, Research Assoc. II

– Dr. Karen Sullivan, Assoc. Librarian

– (tbd) Systems Analyst

US Dept. of Justice, National Institute of Justice

Current Practice

- Almost all digital archiving strategies are predicated on the assumption that data producers will contribute some effort to preparing data for archiving

- Reality: Compliance with requirements (formats, metadata, documentation, etc.) is the rare exception

Current mechanisms for Archive-Ready Data

- Appeals to self-interest
- Appeals to altruism
- Reputation effects
- Service provisions
- Professional norms
- Assistance

Research Design and Methods

- Subjects: Researchers funded by NIJ who are required to deposit data at ICPSR

- Survey and ethnographic analysis of compliance with current deposit guidelines (obstacles, costs, work flow, etc.) (Year 1)

- Laboratory experiments with alternative incentive mechanisms (reputation through citation, scoring rule, enhanced service, formal publication of data and documentation (Years 1and 2)

- Field experiments with promising mechanisms

 (Years 2 and 3)

Anticipated Results

- Carefully designed and field-tested incentive mechanisms that increase cooperation between producers and archives

- Revised data deposit guidelines that incorporate an appreciate of producers' capabilities and motivations

- Greater adherence to data deposit requirements (some of which will be redesigned) that will reduce archiving costs and produce higher quality data