Understanding Qualitative Data Sharing Practices in Social Sciences

Problem: Social Scientists Face Challenges When Sharing Qualitative Data:

The Nature of Qualitative Data
- Qualitative data can be complex, have a loose structure, and may be hard to reuse without a context.

Research Work in Social Science
- Privacy concerns for study participants
- Unclear/hard-to-define data ownership
- Funding and investment in infrastructure
- Data sharing culture

Infrastructural Barriers – Lack of:
- Training, equipment, access to related facilities
- Built infrastructure

Theoretical Foundations for Understanding Data Sharing Practices

+ People (Individuals)
+ Shared norms and value
+ Artifacts
+ Institutions
+ Routines and practices
+ Policies
+ Built technologies

Knowledge Infrastructure (KI) (Edwards et al., 2013)

- Collaboration readiness
- The nature of the work
- Common Ground
- Mgmt., planning, & decision making
- Technology readiness

Theory of Remote Scientific Collaboration (TORSC) (Olson et al., 2008)

Designing the Instrument

Adopting the KI and TORSC, four dimensions are proposed for investigating the determinants of social scientists’ data sharing practices.

- Data characteristics
- Technical Infrastructure
- Organizational context
- Individual motivations

The pre-test questionnaire (on Qualtrics) contained 99 close-ended and three open-ended questions.

Data Collection

57 PhD students and post-docs at Pitt and CMU (USA) completed the survey in December 2015. Participants self-reported their primary research methods:

- Quantitative (N=35)
- Mixed method (N=17)
- Qualitative (N=5)

61% 39%

The top represented disciplines were political science (N=13), education (N=11), economics/business (N=10), and psychology (N=8).

Preliminary analysis

1. The frequency of data sharing are associated with preprint sharing.
2. Highest rated items for (5: strongly agree; 1: strongly disagree)
   - Data sharing conditions – “Altruism”
     I will be willing to share my data “if my data inspire researchers outside of my discipline” (4.16)
   - Perceived benefits - “collaboration opportunities”
     “Sharing data gives me an opportunity to collaborate with other researchers.” (4.07)
3. Participants assessed their institutional supports during the data lifecycle:

   - Data Collection
   - Data Processing
   - Data Analysis
   - Data Sharing

While the existing support of “Data Management” and “Data Analysis” are rated as fairly sufficient, resources for data sharing and reuse are least sufficient.

Next Steps

- Factor analysis and assessment of the reliability among items
- Case study on the Interuniversity Consortium for Political and Social Research (ICPSR):
  - Full-scale national survey of qualitative data sharers
  - Analysis on metadata and metrics of existing shared datasets
  - Interviews with staff and directors to refine the theoretical framework

ACKNOWLEDGEMENTS

We would like to thank Drs. Nora Mattern, Liz Lyon, Sheila Corrall, Jian Qin, Stephen Griffin, and researchers Jessica Borer for their invaluable comments and suggestions on this study. The lead researcher is currently an fellow, and the project is sponsored by Andrew W. Mellon Foundation and guided by the Committee on Coherence at Scale (CoC) and Council on Library and Information Resources (CLIR).

RELATED WORK


Mattern, E., Jeng, W., He, D., Lyon, L., & Brenner, A. (2015). Using participatory design and visual narrative inquiry to investigate the data lifecycle: The top represented disciplines were political science (N=13), education (N=11), economics/business (N=10), and psychology (N=8).

FURTHER READING

- Mattern, E., Jeng, W., He, D., Lyon, L., & Brenner, A. (2015). Using participatory design and visual narrative inquiry to investigate the data lifecycle: The top represented disciplines were political science (N=13), education (N=11), economics/business (N=10), and psychology (N=8).
- Interviews with staff and directors to refine the theoretical framework