E **Evidence Based Library and Information Practice**

Social Scientists' Data Reuse Principally Influenced by Disciplinary Norms, Attitude, and Perceived Effort

A Review of:
Yoon, A. & Kim, Y. (2017). Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories. *Library & Information Science Research*, *39*(3), 224–233. https://doi.org/10.1016/j.lisr.2017.07.008

Reviewed by:
Scott Goldstein
Web Librarian
Appalachian State University Libraries
Boone, North Carolina, United States of America
Email: goldsteinsl@appstate.edu

Abstract

Objective – To propose and test a model grounded in constructs from psychology and information systems to explain data reuse behaviours and practices in the social sciences.

Design – Electronic survey.

Setting – ProQuest's *Community of Science Scholars* database.

Subjects – Included 2,193 randomly selected social scientists associated with U.S. academic institutions.

Methods – An electronic survey was distributed to a random sample of U.S.-based social science scholars from ProQuest's *Community of Science Scholars* database. The survey adapted 21 measurement items for constructs taken from the theory of planned behaviour (TPB) and the technology acceptance model (TAM), including perceived usefulness, perceived effort, and the subjective norm surrounding data reuse.

Main Results – There were 292 valid responses received, giving a response rate of 14.91%. Survey data largely validated the authors' theoretical model. Attitudinal, normative, and resource factors all influence

social scientists' intended data reuse. In particular, perceived usefulness of reusing data and subjective norms surrounding data reuse in one's discipline positively correlate with intentions to reuse data, and perceived concern of reusing data negatively correlate with intentions to reuse data.

Conclusion – Data reuse in the social sciences is influenced by the perceptions and beliefs held by social scientists. Social scientists reuse others' data when they perceive that doing so would improve their research productivity and when their discipline has strong norms of data reuse. They avoid reusing others' data when they believe that doing so is problematic (e.g., if they believe reusing infringes on copyright). Supporters of data sharing, including librarians, are encouraged to apply these findings by proactively educating researchers on the benefits, potential obstacles, and methods of data reuse.

Commentary

This study adds to the literature on data reuse practices in the social sciences. Unlike in the "harder" sciences, social science data may contain qualitative and highly contextual information about human subjects, thereby demanding a higher level of ethical consideration. Previous studies have been primarily exploratory, looking at behaviours and concerns raised by sharing this kind of data. The authors build on this by developing a theoretical model using constructs from the theory of planned behaviour (TPB) from social psychology (Ajzen, 1991) and the technology acceptance model (TAM) from information systems (Davis, 1989).

This evidence summary relies on Glynn's (2006) critical appraisal checklist to determine the validity of the study. A major strength of the study is its data collection methodology. The procedure is fully described and the authors' instrument and data are publically available. Furthermore, the items in the survey were adapted from prior studies and displayed good reliability and convergent and discriminating validity. There are some concerns with the sampling frame, however. The random sample of scholars was obtained from ProQuest's *Pivot* database, which is populated via web harvesting with some unspecified amount of manual correction by *Pivot*'s profile editing team. (This database was formerly referred to as *Community of Science*, and that is the name the authors use throughout the article.) It is not clear if the process used to harvest publicly available profile information introduces any biases into the collection criteria. For example, the demographic breakdown of survey respondents looks as if it might be skewed towards senior, established academics, but no mention of this is made in the text. It is also worth pointing out that 234 invitation emails went undelivered, which is over 10 percent of the total sample. This is a notably high percentage, especially if this is almost entirely due to invalid email addresses, further raising concerns about the original sampling frame.

The study's implications for library and information professionals reinforce what many in the practice are already doing: talking with users about what data is available, addressing copyright and other potential limitations to reusing data, and marketing and providing support for relevant data repositories. The authors suggest research libraries should be more proactive in informing and educating researchers. Librarians may wish to include information on finding and searching data repositories in their instruction, especially in disciplines with a strong norm of data sharing. They may also wish to advocate in favour of open data practices beyond simply what may be required of researchers in some data management plans. Librarians are well-suited to contribute to a culture of data reuse at their institutions.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Process*, *5*0, 179–211.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance in information technology. *MIS Quarterly*, *13*(3), 319–340.

Glynn, L. (2006). A critical appraisal tool for library and information research. *Library Hi Tech*, *24*(3), 387–399. https://doi.org/10.1108/07378830610692154