Exploring Data Curation and Management Programs, Projects, and Services through Metatriangulation

Plato L. Smith II
Florida State University
Communication and Information
Tallahassee, FL 32306, USA

pls08@my.fsu.edu

ABSTRACT
Some data curation, digital curation, and data management programs are developed without working definitions, interdisciplinary research agenda, underlying theoretical frameworks or perspectives, or program evaluations. This paper introduces working definitions, a conceptual framework model, a theory building strategy, assessment, and program evaluation in the development of a proposed data curation and management evaluation framework for contribution to literature on developing data curation and management programs theory development and evaluation.

Categories and Subject Descriptors
D.3.3 [**Models and Principles**]: Systems and Information Theory – *value of information.*

General Terms
Management, Measurement, Human Factors, Theory.

Keywords
Data curation, data management, digital curation, interdisciplinary research, conceptual model, metatriangulation, program evaluation.

1. INTRODUCTION
Literature review shows an interchangeable use of data curation, digital curation, and digital preservation concepts across some institutions developing data curation and digital curation management workshops, certificates, curriculum and professional development programs. Working definitions, theoretical frameworks or perspectives, assessments, standards, funders, institutions, program evaluators' engagement and program evaluations of current programs are needed for baseline and success metrics development of programs attempting to address data curation and management issues across disciplines.

2. PURPOSE
The purpose is to introduce working definitions of data curation, data management, related concepts and an interdisciplinary evaluative framework based on theoretical frameworks for research and development, assessment, and evaluation considerations.

JCDL'12, June 10–14 2012, Washington, DC USA

3. DEFINITIONS
What is a concept? "A concept may be defined as a stable organization in the experience of reality which is achieved through the utilization of rules of relation and to which can be given a name" (Bolton, 1977, p. 23). Concepts are stable structures in the mind created through constant interaction with the environment (Dahlberg, 1978, p. 9) "possessed by specific persons, [which] may be regarded as subjective concepts or 'psychological concepts' (Dahlberg, 1978, p. 10). Persons possessing subjective concepts according to Moravcsik (1977) have six steps of development characterized by having the following abilities: (1) sorting abilities, (2) forming of expectations, (3) ability to handle truth-functional complexities, (4) ability to apply the concept to cases not observable by the senses of either speaker or hearer, (5) having explicit criteria for applying the concept, and (6) being able to consider the concept, characterize it, having a theory about its domain of application, etc. Bolton (1977) and Dahlberg (1978) definitions of concepts may help explore data curation as a concept while Moravcsik (1977) six steps of concepts development articulate the required skills for developing data curation concepts in environments attempting to advance data curation interdisciplinary research.

What is data curation? "Data curation is the active and on-going management of data through its lifecycle of interests and usefulness to scholarship, science, and education. Data curation activities enable data discovery and retrieval, maintain its quality, add value, and provide re-use over time, and this new field includes authentication, archiving, management, preservation, retrieval, and representation. Our program will provide a strong focus on the theory and skills necessary to work directly with academic and industry researchers who need data curation expertise." (GSLIS, 2010).
What is digital curation? "Digital curation is concerned with actively managing data for as long as it continues to be of scholarly, scientific, research and/or administrative interest, with the aim of supporting reproducibility of results, reuse of and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its long-term accessibility and preservation, authenticity and integrity" (DCC, 2010, All Stages section, p. 2).
What is data management? Business dictionary dot com defines data management as "administrative process by which the required data is acquired, validated, stored, protected, and processed, and by which its accessibility, reliability, and timeliness is ensured to satisfy the needs of data users" (Business Dictionary.com, 2012). Data must also be continuously extended, updated, and made secure for reuse through consultation (Otlet, 1903; Bush, 1945; Rayward, 1998) and "discursive formation" (Foucault, 1972) to remain useful to science.

4. RESEARCH ISSUES AND QUESTION

"Good conceptual tools are needed, whether to participate in developments, evaluate their progress, guide researchers to make good use of the results, or critically engage with the very idea of data being usable 'any time, any place'" (Whyte, 2012, p. 205). Inconsistent differentiation of related data curation concepts, definitions, theoretical frameworks or perspectives, standards, best practices in program development and evaluations are some research issues in the data curation and management field. "Problems posed by the increasing volume of [data] and changing scholarly practices (Otlet, 1903) with conflicting, duplicative or erroneous matter" (Rayward, 1998, 28) support the need for development of a conceptual tool that seeks to address these issues.

RQ 1: How can a conceptual framework be used to address some of the data curation issues in the development of data curation programs, projects, and services within and across disciplines?

5. METHODOLOGY

5.1 Multiphase Design

A mixed-methods research design approach utilizing the multiphase design (Creswell & Clark, 2011, p. 70) is conducted. The overall program objective involves an assessment survey (study 1 – qualitative) for assessment evaluation leading to programs selected for evaluation (study 2 - quantitative) leading to evaluated programs (study 3 - mixed methods). Program evaluations include program director interviews, users surveys based on Zinn's Framework (1997), and select standards-based assessment and evaluation tools.

5.2 Completed Tasks

A survey instrument that includes data curation and management, interdisciplinary research, evaluation assessments, and purposive sampling questions is completed. A proposed data curation and management evaluation framework research design (see Fig. 2.) has also been developed to assist in guiding the research which is based on a conceptual model framework (See Fig. 1.). The survey instrument and framework have been shared with major professor for scholarly rigor and a funder for feedback, edits, and future implementation.

6. CONCEPTUAL MODEL FRAMEWORK

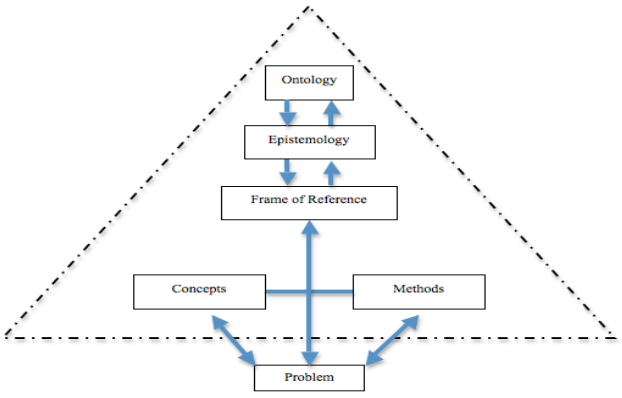

Figure 1. Conceptual Model Framework (Solem, 1993, p. 595)

What is a conceptual framework? "Conceptual frameworks are descriptive categories placed within a broad structure of both explicit and assumed propositions. The theory of symbolic interaction is best termed a conceptual framework. Concepts such as "self" [actor], "object" [data], "act" [curation], "interaction" [lifecycle] and "socialization" are used to analyze the data, and a systematic image of the empirical world is assumed" (Denzin, 2009, p. 67).

What is interdisciplinary research? "Interdisciplinary research is a mode of research by teams or individuals that integrates information, data, techniques, tools, perspectives, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice" (Committee on Facilitating Interdisciplinary Research, Committee on Science, Engineering, and Public Policy, 2004, p. 2).

What is "discursive formation"? "Discursive formation involves the promulgation of new ideas, identification of new phenomena, changes in language practices, and elaboration of new terminology. Discursive formation requires the creation of new formal structures of communicating individuals and the development of new tools and techniques for information handling" (Rayward, 1998, p. 22).

6.1 Metatriangulation – Theory Building

Metatriangulation (Lewis & Grimes, 1999) is a process of building theory across multiple paradigms (Gioia & Pitre, 1990). Metatriangulation conceptualizes the process of theoretical triangulation (Denzin, 1970). In situations where theories are not readily available or insufficiently useful to guide research, one can take the steps of using "metatriangulation as a conceptualized theoretical triangulation (Denzin, 1978) process of building theory from multiple paradigms analogous to traditional single-paradigm triangulation" (Gioia & Pitre, 1990; Lewis & Grimes, 1999, p. 676). "It is conventionally assumed that triangulation is the set of multiple methods in the study of the same object" (Campbell & Fiske, 1959; Webb, 1966; Denzin, 2009, p. 301).

"Focusing on discipline-specific methods is both too narrow and too transitory given the transformative influence of rapidly changing technologies" (National Academy of Sciences, National Academy of Engineering, & Institute of Medicine, 2009, p. xi) and how data management and curation programs, projects, and services are taking place across disciplinary and geographic boundaries (Star & Ruhleder, 1996). "Metatriangulation is particularly useful on topics where a rich repository of published research with contradictory, confusing, and fragmented results exist" (Saunders et al., 2003; Madlberger & Roztocki, 2010, p. 2).

6.1.1 Multiparadigm reviews

"Multiparadgim reviews involve recognition of divides and bridges in existing theory to reveal the impact of theorists' underlying, and often taken-for-granted, assumptions on their understandings of organizational phenomena.

6.1.1.1 Multiparadigm research

Multiparadigm research applies divergent lenses empirically, which includes parallel and sequential studies.

6.1.1.2 Multiparadigm theory building

Multiparadigm theory building manages bounded rationality and accommodate opposing views within a metaparadigm perspective" (Lewis & Grimes, 1999, pp. 673-675).

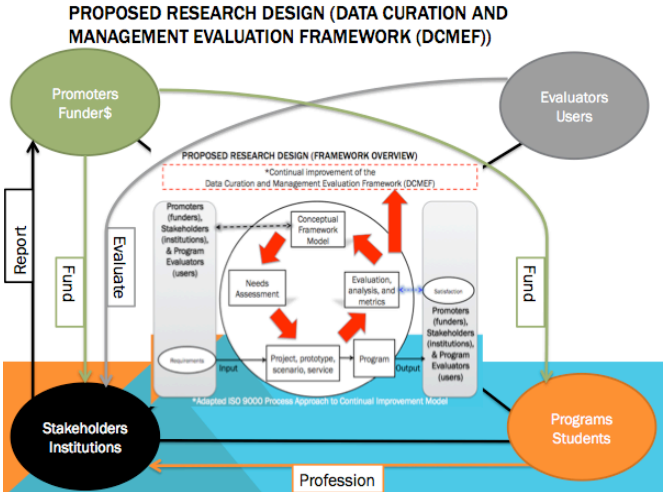

Figure 2. Data Curation and Management Evaluation Framework (DCMEF)

7. ORIGINALITY

This research project will explore how a conceptual model framework and metatriangulation may be enabled within an evaluation framework to further interdisciplinary research in data curation and management programs, projects, and services for improved funders, institutions, and program evaluators support, collaboration, and engagement. Projected project completion is late Fall 2012.

8. ACKNOWLEDGMENTS

Thanks to ACM SIGCHI for allowing me to modify templates they had developed and Dr. Paul Marty, Dr. Corrine Jörgensen, and Dr. Melissa Johnston for scholarly guidance and feedback.

9. REFERENCES

[1] Bolton, N. (1977) Concept formation. Pergamon Press.

[2] Bush, V. (1945). As we may think. Retrieved May 27, 2012 from As We May Think.

[3] Business Dictionary.com. (2012). Data management. Retrieved May 27, 2012 from Data management.

[4] Committee on Facilitating Interdisciplinary Research, Committee on Science, Engineering, and Public Policy. (2004). Facilitating interdisciplinary research. National Academies. Washington: National Academy Press, p. 2.

[5] Creswell, J. & Clark, V. (2011). Designing and conducting mixed methods research. Sage Publications, Inc. Thousands Oaks, California.

[6] Dalhberg, I. (1978). Ontical structures and universal classification. Sara Ranganthan Endowment for Library Science, Bangolore.

[7] DCC. (2010). DCC Curation 101 Manual. All stages section.

[8] Denzin, N. (2009). The research act: a theoretical introduction to sociological methods. Piscataway, NJ: Transaction Publishers, Rutgers-The State University of New Jersey.

[9] Denzin, N. (1978). The research act. New York: Wiley.

[10] Fieldling, N. G. & Fieldling, J. L. (1986). Linking data. Beverly Hills, CA: Sage.

[11] Foucault, M. (1972). The archeology of knowledge and the discourse on language. New York: Pantheon Books.

[12] Gioia, D. & Pitre, E. (1990). Multiparadigm perspectives on theory building. Academy of Management Review, 15(4): 584-602.

[13] GSLIS. (2010). Graduate school of library and information science. The iSchool at Illinois. Data Curation. Retrieved May 27, 2012 from Data Curation.

[14] Lewis, M., & Grimes, A. (1999). Metatriangulation: building theory from multiple paradigms. Academy of Management Review, 24(4): 672-690.

[15] Madlberger, M., & Roztocki, N. (2010). Digital cross-organization collaboration: a metatriangulation review. Systems Sciences, (HICSS), 2010 43rd Hawaii International Conference on, p. 1-10. doi: 10.1109/HICSS.2010.149.

[16] Otlet, P. (1903). International organisation and dissemination of knowledge: Selected essays of Paul Otlet. Translated and Edited with an Introcution by W. Boyd Rayward. FID Publication 684; Amsterdam: Elsevier.

[17] Rayward, W. B. (1998). The origins of Information Science and the International Institute of Bibliography/International Federation for Information and Documentation (FID). In. T Hahn and M. Buckland (Eds.), Historical Studies in Information Science (pp. 22-33).

[18] Saunders, C., Carte, T. Jasperson, J., & Butler, B. (2003). Lessons learned from the trenches of metatriangulation research. Communications of Association for Information Systems, 11, pp. 245-270.

[19] Solem, O. (1993). Integrating scientific disciplines: an evergreen challenge to systems science. In F. Stowell, D. West, & J. Howell (Eds.), Systems Science: Addressing Global Issues (pp. 593-598). New York: Plenum Press.

[20] Star, S. & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: design and access for large information spaces. *Information Systems Research*, 7(1): 111-134.

[21] Webb, E. (1966). Unconventionality, triangulation, and inference. In the Proceedings of the 1966 Invitational Conference on Testing Problems, pp. 34-43. Princeton, NJ: Educational Testing Service.

[22] Whyte, A. (2012). Emerging infrastructure and services for research data management and curation in the UK and Europe. In G. Pryor (Ed.). Managing Research Data. (pp. 205-234). London: Facet Publishing.

[23] Zinn, L. (1997). Supports and barriers to teacher leadership: reports of teacher leaders. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, March 24-28. 43 pages.

[24] Rossi, P., Lipsey M., & Freeman, H. (2004). Evaluation. A systematic approach. 7th ed. Thousands Oaks: Sage