

# Fifteenth International Conference on Grey Literature - The Grey Audit: A Field Assessment in Grey Literature



**Data management & curation services (DMC): exploring  
stakeholders opinions**



Plato L. Smith II, Doctoral Candidate,  
pls08@my.fsu.edu  
Florida State University  
College of Communication and Information, USA  
GL15 - Bratislava, Slovak Republic - December 3,  
2013

# Outline



- ① Introduction
- ② Data Management and Curation (DMC)
- ③ Statement of Problem
- ④ Research Questions
- ⑤ Methodology
- ⑥ DMC Survey - FSU IRB HSC# 2012.9198
- ⑦ Limitations and Significance of Study
- ⑧ References

# Introduction

- Guidelines
- Policies
- Workflows
- Frameworks/Models

Data  
Management  
Planning

- Research
- Data creation
- Representation
- Publication

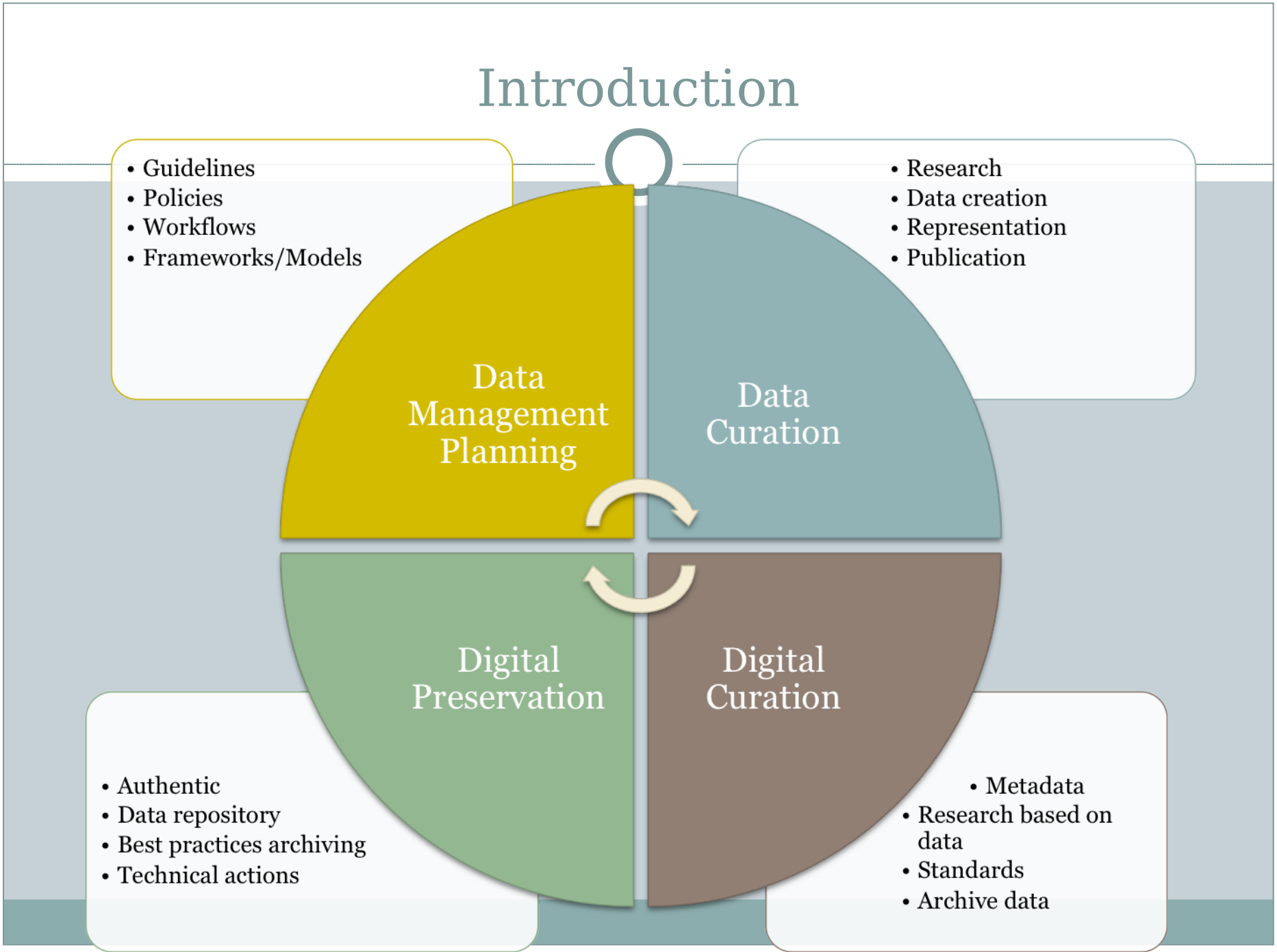
Data  
Curation

Digital  
Preservation

- Authentic
- Data repository
- Best practices archiving
- Technical actions

Digital  
Curation

- Metadata
- Research based on data
- Standards
- Archive data



# Data Management and Curation (DMC)



## Four Key Concepts of Data Management & Curation:

1. Data Management Planning (Entire data lifecycle)
2. Data Curation (Level 1 Curation - Traditional academic information flow)
3. Digital Curation (Level 2 Curation - Information flow with data archiving)

4. Digital Preservation (Level 3 Curation)

*"A record if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted." - Vannevar Bush, 1945*

# Data Management and Curation (DMC)



DMC practices include four major data lifecycle management processes that:

1. Fulfill departmental, institutional, organizational policies & data management requirements;
2. Provide data creation (primary, secondary, tertiary data), data publication, minimal data description;
3. Facilitate added value (metadata), management & storage of archived data over data lifecycle;
4. Integrate a series of technical & strategic actions and consultations to ensure continual data archiving, authenticity, integrity, and stewardship (Lord, 2003; Pennock, 2006; DCC, 2007).

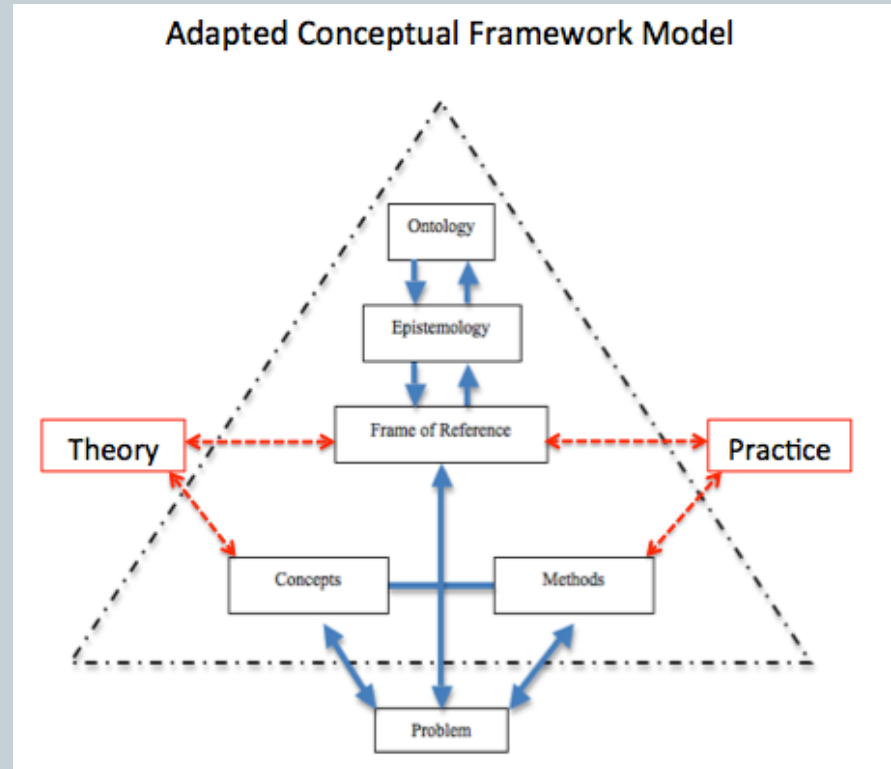
# Statement of Problem



- Definitional confusion of DM key concepts
- Competing models/frameworks fragmentation
- Undeveloped theory of digital preservation and theory of digital curation
- Under-utilization of relevant standards, best practices, and guidelines where appropriate
- A need to improve DMC within & across disciplines
- DMC activities and practices vary across disciplines
- Multiple disciplines face massive data storage issues
- It is “impossible to define all the terms of one theory in the vocabulary of the other” (Kuhn, 1982, p. 669).

# Research Questions

1. How can definitional confusion of key DMC concepts be resolved within and across disciplines?
2. What are some of the theoretical frameworks used to address data management and curation issues?
3. Can multiple paradigm perspectives help develop DMC theory?



(Burrell & Morgan, 1979;  
Morgan & Smircich, 1980;  
Morgan, 1983; Solem, 1993)

# Methodology



- ❖ Design – Quantitative (online survey) research method
  - Qualtrics online survey – 10 questions
  - Primary survey questions focus on (1) DMC key concepts, (2) theoretical frameworks/perspectives, (3) elements of data management plan, and (4) data seal of approval assessment guidelines
  
- ❖ Participant Selection – Professionals/researchers affiliated with data management and curation (DMC)
  - Professionals from formal and informal networking contacts (i.e. professional list serves, conferences, & workshops)
  - n = 64 (64 starts & 53 completes: 83% completion rate)
  
- ❖ Sampling – Funders (promoters), stakeholders (institutions), & users from the US and foreign countries
  - Senior management, deans, faculty, funding program officers, researchers, scientists, practitioners, librarians, publisher, consultant, commercial





# DMC Survey - FSU IRB HSC# 2012.9198

- ✓ Approved 11/2/2012
- ✓ Launched 11/5/2012
- ✓ Closed 12/5/2012
- 12 questions - online
- 83% completion rate
- Diverse stakeholders
- Multiple disciplines
- Various perspectives
- Cultural opinions
- ◆ Top 5 for Q7 - Q9

## Q6. Key Concepts

Data curation, digital curation, and digital preservation are independent yet interrelated concepts. 80% agree (45 out 56).

## Q7. Theoretic Frameworks

1. Pragmatism (67%)
2. Ethnography (64%)
3. Grounded Theory (48%)
4. Autoethnography (36%)
5. Phenomenology (33%)

### Data Management & Curation

## Q8. Elements of Data Management Plan

1. IP Rights (98%)
2. Format (94%)
3. Metadata (94%)
4. Storage and backup (94%)
5. Archiving/preservation (94%)

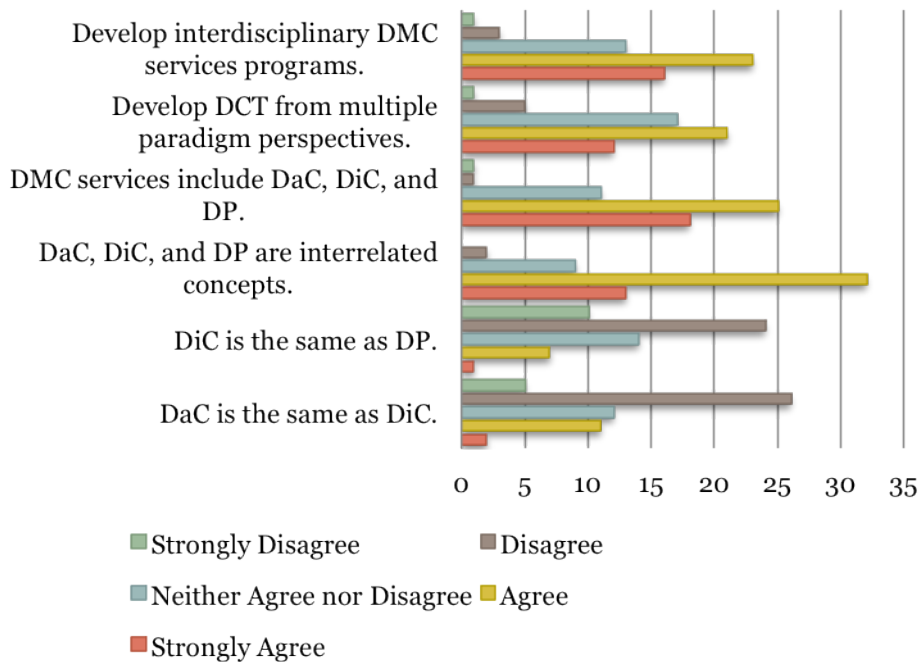
## Q9. Data Seal of Approval Assessment Guidelines

1. Guideline #7 (94%)
2. Guideline #1 (90%)
3. Guideline #6 (88%)
4. Guideline #3 (86%)
5. Guideline #9 (84%)

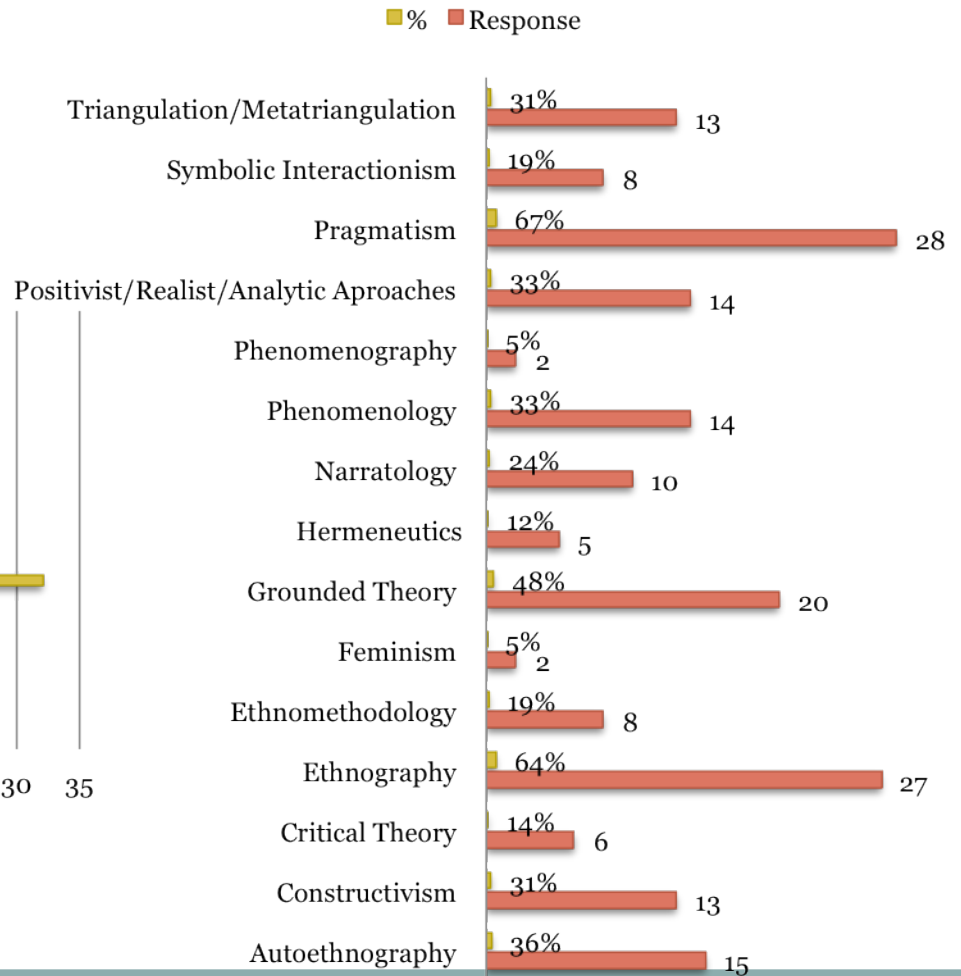
# DMC Key Concepts and Theoretical Frameworks

DMC = Data management and curation  
 DCT = Data Curation Theory  
 DaC = Data curation  
 DiC = Digital curation  
 DP = Digital preservation

## DMC Key Concepts



## Theoretical Frameworks



# Limitations and Significance of Study



## Limitations of Study

- Population selection bias
- Small sample size
- Survey questions bias
- Assumption that participants are familiar with DMC key concepts
- Partial, incomplete, & drop out survey responses
- The study may lack transferability & generalizability.

## Significance of Study

- The study articulates the differentiation of key DMC concepts for definition clarification, linking, and concept mapping.
- The study applies theoretical and practical knowledge to underdeveloped research on theory of data management and curation (DMC).
- The study may help in DMC theory development within & across disciplines.

# References



- Bodner, G. M. & Orgill, M. (2007). Theoretical frameworks for research in chemistry/science education. Pearson Education, Inc.: Upper Saddle River, NJ.
- Burrell, G. & Morgan, G. (1979). Sociological paradigms and organizational analysis. London: Heineman.
- DCC. (2007). The DCC Curation Lifecycle Model. 3<sup>rd</sup> International Digital Curation Conference.
- Gioia, D. & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management Review*, 15(4): 584-602.
- Kuhn, T. (1982). Commensurability, Comparability, Communicability. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 1982, Volume Two: Symposia and Invited Papers (1982), pp. 669-688.
- Lewis, M., & Grimes, A. (1999). Metatriangulation: building theory from multiple paradigms. *Academy of Management Review*, 24(4): 672-690.
- Lord, P. & Macdonald, A. (2003). E-Science Curation Report: data curation for e-Science in the UK: an audit to establish requirements for future curation and provision.
- Patton, M. Q. (2002). Qualitative research and evaluation (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- Pennock, M. (2006). JISC Digital Preservation: continued access to authentic digital assets briefing paper.
- 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.