Sharing is Caring: Challenges and Opportunities in Producing Open Data

Overview and Introduction, Hailey Mooney
Ecology Perspective, Patricia Soranno
Geography Perspective, Joe Messina
Archaeology Perspective, Lynne Goldstein
Research Integrity Considerations, James Pivarnik
Intellectual Property Considerations, Connie Berg

MSU CyberInfrastructure Forum – October 22, 2015

Overview and Introduction:
Scholarly Communication and Data

Hailey Mooney
Data Services Coordinator
and Social Sciences Librarian
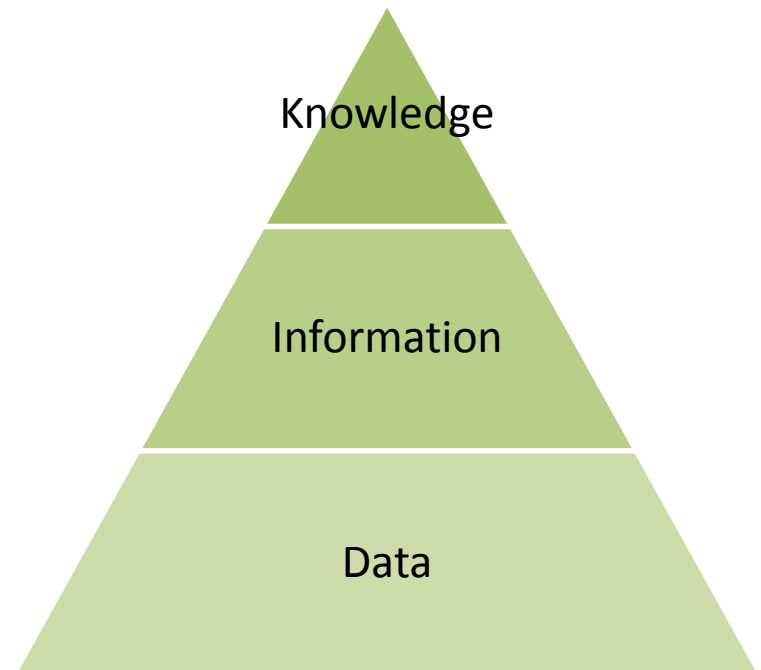
Michigan State University
LIBRARIES

- The nature of scientific knowledge and data

- Open data and the forces behind data sharing and publication

- Data sharing and publication practices

Fundamentals of Scholarly Communication
Production of Knowledge and the Scientific Ethos

- Universalism
- Communism
- Disinterestedness
- Organized Skepticism

Information

Robert K. Merton
The Sociology of Science
Theoretical and Empirical Investigations
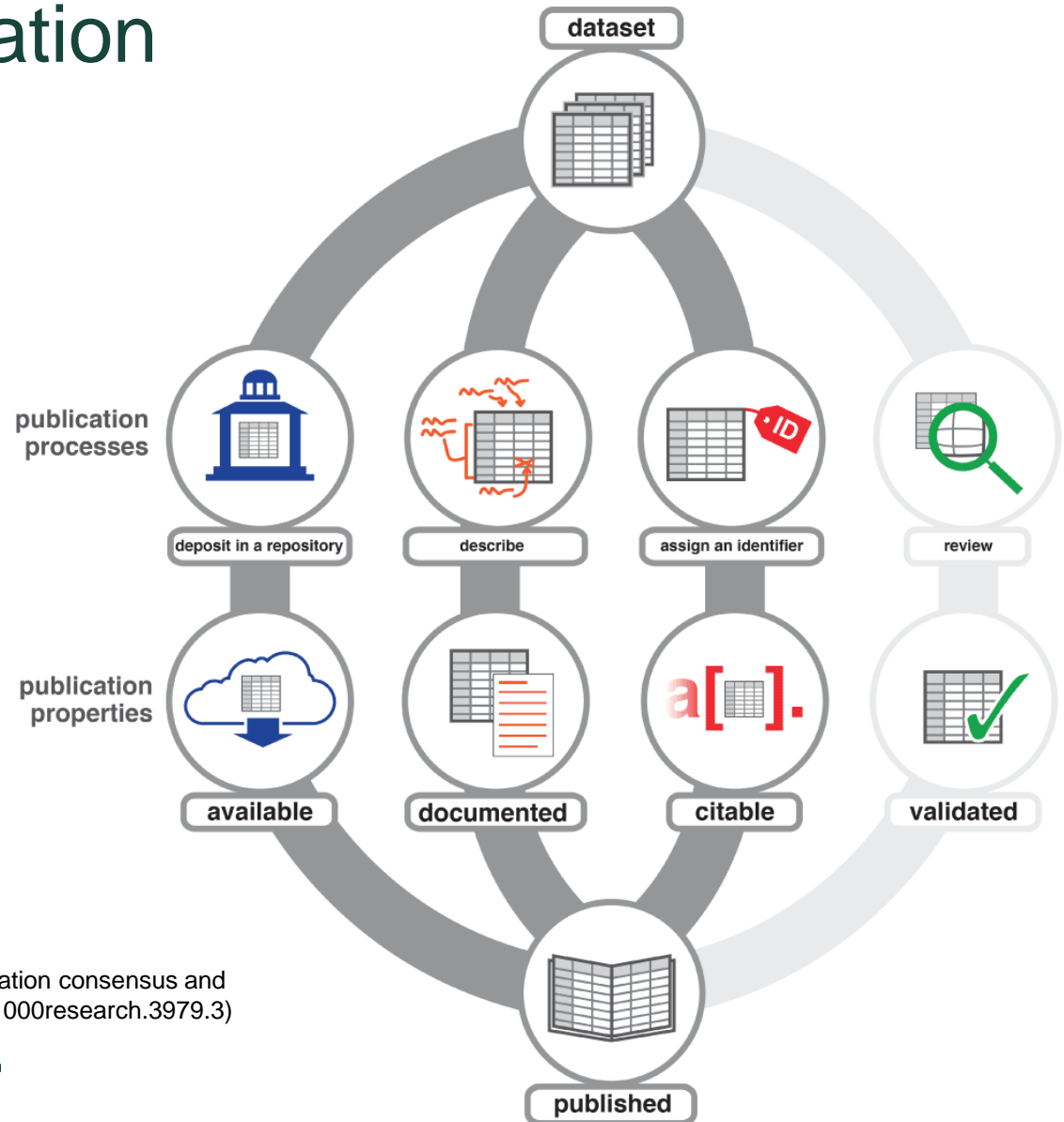Edited and with an Introduction by Norman W. Storer

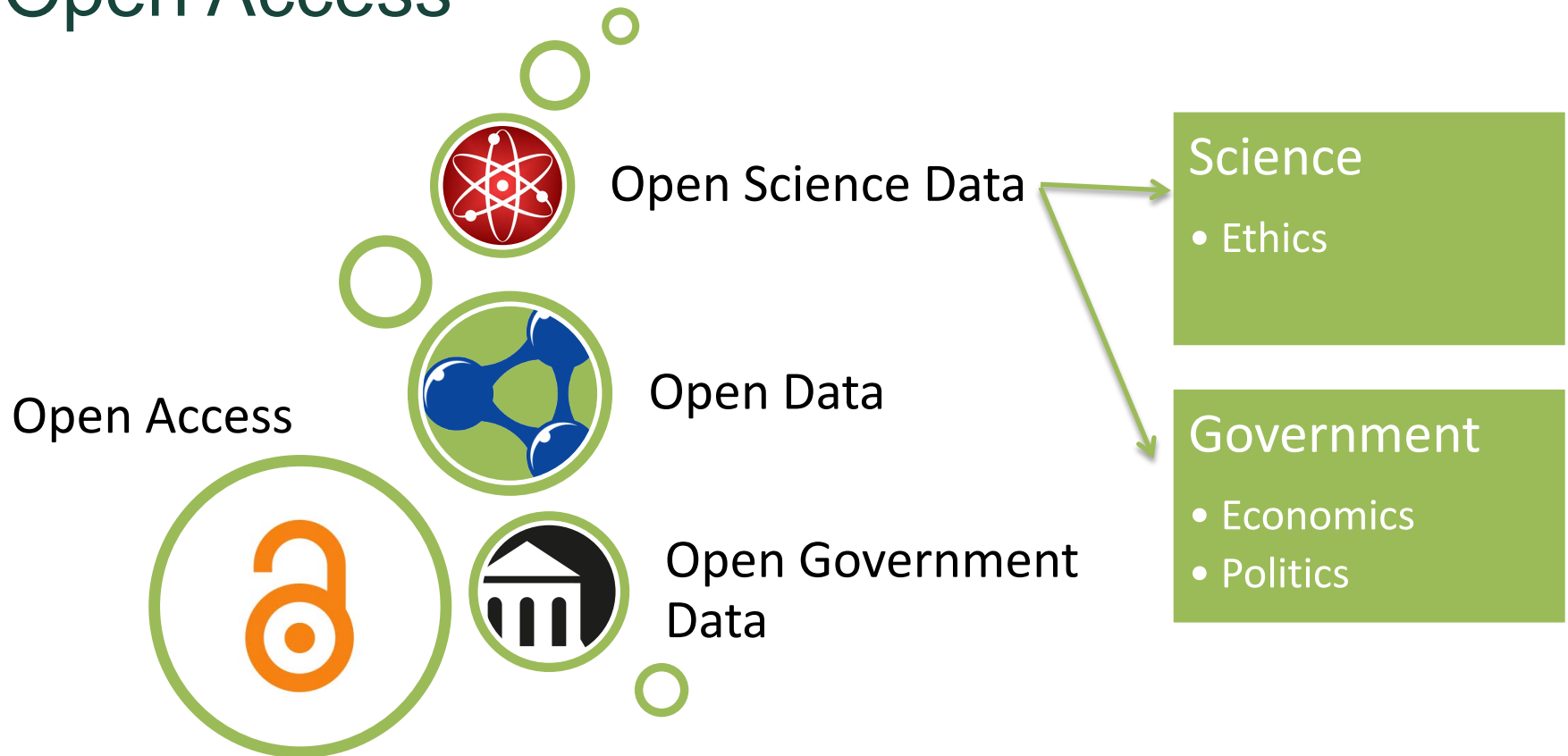# Data?

Data as Publication

Figure 1. To be published, datasets are typically deposited in a repository to make them available, documented to support reproduction and reuse, and assigned an identifier to facilitate citation.

Kratz J and Strasser C (2014) [version 3] Data publication consensus and controversies. *F1000Research* 3:94 (doi: 10.12688/f1000research.3979.3)

Federally funded research data

1980
Forsham v. Harris

- Research data not subject to FOIA

2003
NIH Data Sharing Plans

- Grants over $500,000 require plans

2011
NSF Data Management Plans

- Grants require supplementary DMP

1999
Data Access Act of 1999

- OMB Circular A-110 revised: data produced with Federal monies subject to FOIA

2010
America Competes Reauthorization Act

- OSTP must coordinate policies for dissemination and stewardship of scholarly publications and data produced with Federal funds

2013
OSTP memo: Increasing Access to the Results of Federally Funded Research

- Requires agencies to develop plans for public access to publications and data

Fischer, E. A. (2013). *Public Access to Data from Federally Funded Research: Provisions in OMB Circular A-110* (Congressional Research Service). Retrieved from HTTP://congressional.proquest.com.proxy2.cl.msu.edu/congressional/docview/t21.d22.crs-2013-rsi-0116?accountid=12598

**American
Sociological
Association**

Code of Ethics

13.05 Data Sharing

(a) Sociologists share data and pertinent documentation as a regular practice. Sociologists make their data available after completion of the project or its major publications, except where proprietary agreements with employers, contractors, or clients preclude such accessibility or when it is impossible to share data and protect the confidentiality of the data or the anonymity of research participants (e.g., raw field notes or detailed information from ethnographic interviews).

(b) Sociologists anticipate data sharing as an integral part of a research plan whenever data sharing is feasible.

(c) Sociologists share data in a form that is consonant with research participants' interests and protect the confidentiality of the information they have been given. They maintain the confidentiality of data, whether legally required or not; remove personal identifiers before data are shared; and, if necessary, use other disclosure avoidance techniques.

(d) Sociologists who do not otherwise place data in public archives keep data available and retain documentation relating to the research for a reasonable period of time after publication or dissemination of results.

(e) Sociologists may ask persons who request their data for further analysis to bear the associated incremental costs, if necessary.

(f) Sociologists who use data from others for further analyses explicitly acknowledge the contribution of the initial researchers.

Data sharing realities

Table 11. Data sharing.

	Agree Strongly	Agree Somewhat	Neither Agree Nor Disagree	Disagree Somewhat	Disagree Strongly
I share my data with others.	418 (32.3%)	551 (42.6%)	199 (15.4%)	95 (7.3%)	30 (2.3%)
Others can access my data easily.	150 (11.6%)	317 (24.6%)	310 (24%)	307 (23.8%)	207 (16%)

doi:10.1371/journal.pone.0021101.t011

Tenopir C, Allard S, Douglass K, Aydinoglu AU, et al. (2011) Data Sharing by Scientists: Practices and Perceptions. PLoS ONE 6(6): e21101. doi:10.1371/journal.pone.0021101
http://www.plosone.org/article/info:doi/10.1371/journal.pone.0021101

Lack of data sharing and publication

- Data sharing norms = informal
- Inadequate support for data management
- Inadequate infrastructure and publication mechanisms
- Legal and ethical concerns

Digital Repositories

- Institutional
- Data archive

- Private data repository service

Journals

- Supplementary Materials
  - "A Pandora's Box of Issues Needing Best Practices"

- Data Paper
  - "Data publication by proxy"

Carpenter, T. (2013). Standards Column - Journal Article Supplementary Materials: A Pandora's Box of Issues Needing Best Practices. *Against the Grain*, *21*(6). Retrieved from http://docs.lib.purdue.edu/atg/vol21/iss6/6

Kratz J and Strasser C (2014) [version 3] Data publication consensus and controversies. *F1000Research* 3:94 (doi: 10.12688/f1000research.3979.3)

Citation allows Priority Claims

- Public-private paradox
  - Sharing establishes ideas as private property via the citation

Stephan, P. E. (2004). Robert K. Merton's perspective on priority and the provision of the public good knowledge. *Scientometrics*, *60*(1), 81–87. http://doi.org/10.1023/B:SCIE.0000027311.17226.70

Libraries are your partner in the research enterprise

- Data management
  - Grant proposals
  - Best practices
- Trustworthy data publication venues

Michigan State University
LIBRARIES