February 2010

An Organizational Context for Scientific Data Practices

Kimberly L. Douglass
University of Tennessee - Knoxville, kdougla2@utk.edu

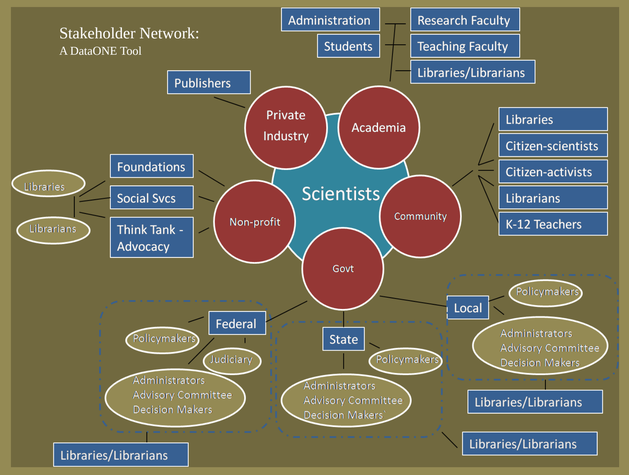
An Organizational Context for Scientific Data Practices

(Kimberly Douglass, Lei Wu, Carol Tenopir, Suzie Allard, Maribeth Manoff, Eleanor Read) University of Tennessee, Bruce Wilson[ORNL], Patricia Cruse California Digital Library, Mike Frame U.S. Geological Survey

Research Question

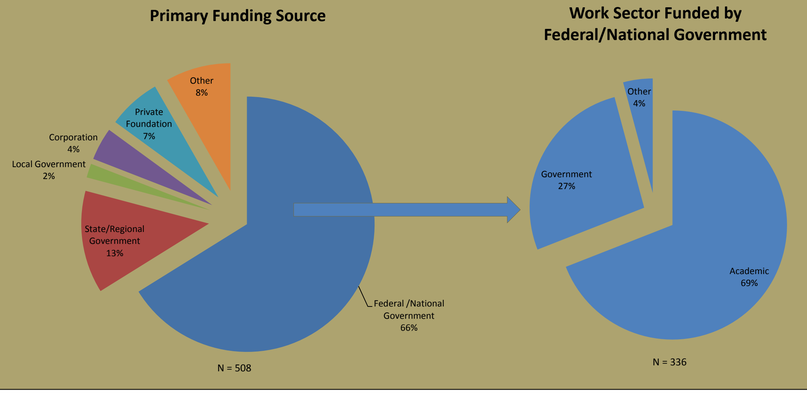
How are government-housed and government-funded scientists constrained in data sharing?

DataONE



- NSF-funded project
- Enables science
- Universal access to data about life on Earth and the environment that sustains it

Demographics – Survey of Data Practices



Preliminary Findings

If some or all of your data are available to others, these data are available on my organization' website.	**Academia 52.5%** **Government 75.9%**
*Pearson's R (Value = 31.375 – df =12 - .significance =.002)*** *Phi (Value=.332 - significance=.002)***	

My organization or project has a formal established process for storing data beyond the life of the project (long-term).	**Academia 39.3%** **Government 57.3%**
*Pearson's R (Value = 27.450 – df =16 - .significance = .037)** *Phi (Value=.288- significance=.037)**	

Tell us how much you agree with the statement, "my organization or project provides training on best practices for data management".	**Academia 20.6%** **Government 24.5%**
*Pearson's R (Value = 32.731 - df = 16 - significance =.008) *** *Phi (Value =.314 - significance =.008)***	

I would be more likely to make my data available if I could place conditions on access.	**Academia 67.1%** **Government 61.8%**
*Pearson's R (Value =31.126 – df =16 - significance=.013)** *Phi (Value=.309 - significance =.013)**	

If your data are not available electronically to others, why not (check all that apply)?

Insufficient Time	Lack of Funding	Do not Have Rights to Make Data Public	No place to Put Data
Academia 57.8% Government 58.3%	Academia 37.7% Government 50.0%	Academia 20.6% Government 27.8%	Academia 21.1% Government 20.8%

Analysis

- More data available on government websites
- Storage is more of an issue in academia.
- Training on best practices for data management is an issue for both sectors.
- Both would make data available conditionally.
- Time and funding are issues for both sectors.

What Have We Learned?

Government-housed and government-funded scientists differ only in information available on their websites and processes for data storage.

Significant at the .05 level, significant at the .010 level**, significant at the .001 level****

Acknowledgements

Support for DataONE is provided through NSF DataNET-0830944 Cooperative Agreement. The principal investigator of DataONE is William Michener, University of New Mexico.