

DataONE: A Glimpse into the Practices of Data Managers

Eleanor J. Read
Usability & Assessment Working Group

eread@utk.edu

IASSIST 2012

DataONE: Data Observation Network for Earth

- www.dataone.org
- Funded by the National Science Foundation
- Building cyberinfrastructure to ensure the preservation of and access to environmental science data. There is also a focus on community engagement and outreach efforts related to DataONE.

Data Managers Survey

- Created by members of the Usability & Assessment Working Group led by Carol Tenopir (University of Tennessee School of Information Sciences)
- Distributed to a variety of groups and individuals by WG members and others
- Sent to IASSIST listserv
- 80 respondents

Respondent Characteristics

- Primary work sector:
 - Government: 54/68%
 - State/Provincial/Tribal: 7/13%
 - Federal/National: 47/87%
 - Academic: 22/28%
 - Non-profit: 2/2%
 - Other: 2/2%

Respondent Characteristics

- Primary subject discipline of the data:
 - Environmental science: 46 responses
 - Ecology: 37
 - Biology: 36
 - Physical sciences: 30
 - Hydrology, social sciences: 29
 - Geology: 25
 - Atmospheric Science: 23

Respondent Characteristics

- Primary types of data managed:
 - Observational (no manipulation): 47 responses
 - Data models: 31
 - Biotic surveys: 29
 - Experimental (some degree of manipulation): 26
 - Remote-sensed abiotic data: 25
 - Social science survey: 22
 - Remote-sensed biotic data, abiotic surveys: 21
 - Interviews: 13

Providing Data Management

- How many researchers do you support?
 - 0 3(4%)
 - 1-5 7(9%)
 - 6-10 4(5%)
 - 11-20 3(4%)
 - 21-30 4(5%)
 - 31-40 1(1%)
 - 41-50 3(4%)
 - More than 50 49(66%)
- What percent of the data your agency generates is managed by you?
 - 1-24% 42(67%)
 - 25-49% 8(13%)
 - 50-74% 3(5%)
 - 75-100% 10(16%)

If your data are not available electronically to others...Why?

- Data are restricted: 29 responses
- No administrative incentives/requirements: 23
- P.I. doesn't want the data available: 22
- Lack of staff: 21
- Lack of time: 20
- Lack of funding: 17
- Lack of technical infrastructure: 16
- Lack of standards, insufficient time to make available: 15
- Data may be temporarily embargoed: 14

Increasing Data Deposit / Storage

- Increase funds to manage data: 44
- Data deposit required/Extra Assistance: 29
- Credit: 28
- Training: 26
- Support incremental data deposit: 16

Data Management Plans

- Have a data management plan for data that you manage?
- Yes: 49/83% ; No: 10/17%
- Confident the DMP will ensure long-term access to the data:
- Agree: 73%, Disagree: 19%

Data Management Plans

- My DMP is regularly reviewed and updated: Agree: 56%, Disagree: 29%
- Researchers I work with understand the DMP: Agree: 37%, Disagree: 29%

Conditions of Data Sharing

Important or Very Important

- Formal acknowledgement of the data providers and/or funding agencies in all disseminated work: 95%
- Formal citation of the data providers and/or funding agencies in all disseminated work: 99%
- Prior approval before results based on the data are disseminated in any format: 41%

Conditions of Data Sharing

Important or Very Important

- At least part of the costs of data acquisition or retrieval must be recovered: 20%
- Reprints of articles that make use of the data must be provided to the data provider: 47%

Conditions of Data Sharing

Important or Very Important

- The data provider is given a list of all products that make use of the data: 58%
- Permission for use is obtained: 42%

Questions?