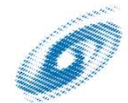


# The Promise? – *or* – Promise! of Data in eResearch

Many challenges, multiple  
solutions, diverse outcomes

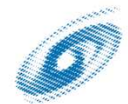
eResearch08, Oxford  
September 12, 2008

Ann Zimmerman  
School of Information  
University of Michigan



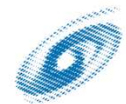
# Key points

- The way in which digital databases are created affects practice, attitudes, & relationships.
- In order to assess or predict outcomes of a particular approach, we must define what we are trying to achieve.



# Driving questions

- How do the origins of digital databases affect
  - research practice,
  - researchers' attitudes toward data sharing,
  - relationships between researchers and others such as data managers, and
  - reusability of data?



# Approaches to data sharing

## Challenges:

- Heterogeneity of all kinds (e.g., semantic, syntactic, methodological)
- Lack of reward for sharing; lack of recognition for reuse
- Legal & economic



**Sharing is considered before data are collected & solutions are developed to enhance data reuse.**

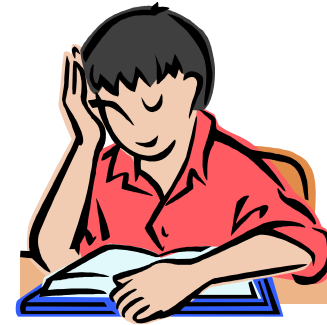
**Hybrid approaches**

**Data sharing is handled “after the fact” by people other than those who produced the data.**

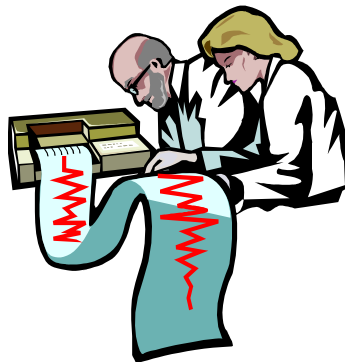
# Wormbase: Sharing “after the fact”



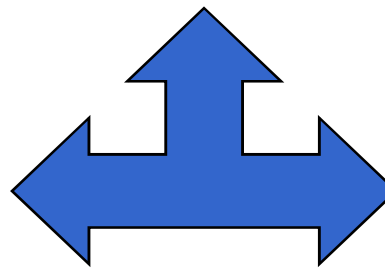
Step 1: Scientist publishes a paper in a journal.



Step 2: A data curator scans journals for data, extracts data and descriptive information, and enters it all into a computer.

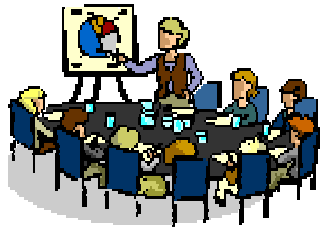


Step 4: Other people use the data. They provide input that results in additions of new types of data or corrections to the database.



Step 3: Data are integrated with other data into one database and made available to anyone via the Internet.

# AfCS: Reuse considerations part of research



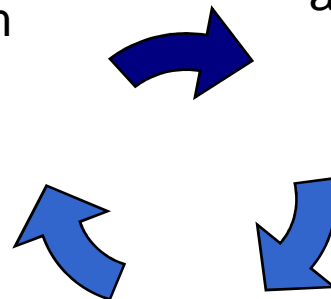
Step 1: Individuals get together to decide upon standard data collection protocols. Team includes data experts.



Step 2: Every laboratory uses the agreed upon methods to collect data.



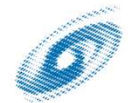
Step 4: Other people use the data. They provide input that influences how new data are collected or that result in corrections to existing data.



Step 3: Data are integrated and made available to anyone via the Internet.

# What is the goal of data sharing?

- Is it to...
  - influence research practice;
  - relieve researchers of the burdens of data sharing; or
  - bring the expertise of others more directly into the research process?



# Acknowledgments

- Science of Collaboratories project: Gary Olson, PI; Judy Olson, & Nathan Bos
- National Science Foundation Grant IIS 0085981

asz@umich.edu

