UNDERSTANDING PERSPECTIVES ON REUSING NEUTRON DATA AT OAK RIDGE NATIONAL LABORATORY

DEVAN RAY DONALDSON, PH.D., AND SHAWN MARTIN

Email: drdonald@indiana.edu
@hoosierdevan
WHO AM I?

Devan Ray Donaldson
Assistant Professor of Information Science
Department of Information and Library Science
School of Informatics and Computing
Indiana University, Bloomington
Ph.D. in Information Science, University of Michigan
RDA US Data Share Fellow
WHY AM I HERE?

Because of Thomas Proffen
Because Frank asked me to come speak :0)
To understand perspectives on data sharing in a field that has traditionally focused more on sustaining use of data by those who created them as opposed to enabling reuse of data by others.
STUDY DETAILS

Focus groups with:
Data consumers (n=3)
Data managers (n=5)
Data producers (n=5)
Data producers generate raw data (unprocessed numbers and descriptions) from which they can construct reduced data (a set with extraneous data removed and more complete descriptions).

Data managers then produce reduced data that harmonizes the unprocessed data with theoretical models which can be used to create modeled data, or data that demonstrates how results do/do not conform to theoretical models.

Data consumers utilize modeled data to create research and scholarship demonstrating how materials function on an atomic level.
FINDINGS: DATA CONSUMERS

1. Identified reasons for reusing data
2. Discussed information they needed to know about data
3. Articulated the importance of journal articles
4. Described barriers to reuse
5. Expressed a desire for discoverability
PARTICIPANT CHARACTERISTICS

- Expressed interest in data reuse
- 2 research scientists; 1 professor
- Interests: theory of magnetism, condensed/soft matter physics
- Multiple years of experience with neutron data and Oak Ridge facilities
PARTICIPANT CHARACTERISTICS

- Expressed interest in data reuse
- 2 research scientists; 1 professor
- Interests: theory of magnetism, condensed/soft matter physics
- Multiple years of experience with neutron data and Oak Ridge facilities
REASONS FOR REUSING

-To compare/verify a result against their own measurements
-To test a new theory using existing data
WHAT REUSERS NEED TO KNOW

1) How the data were produced
2) How the sample was prepared
3) What the units of measurement are
4) How the temperature was determined
1) Journal articles provide context for data
2) Participants articulated interest in reproducing charts and graphs
BARRIERS TO REUSE

Technical barriers:
  e.g., Lack of expertise in software
Consumers of neutron data want to know:
1. What other measurements have been created for particular problems
2. Particular characteristics across data sets (e.g., temperature readings)
RECOMMENDATIONS

Policy recommendations
Technical recommendations
Provide Principal Investigators with the option to make their data accessible and openly available if they choose.
SYSTEM RECOMMENDATIONS

1) Include metadata about how the data were produced, how the sample was prepared, what the units of measurement are, and how the temperature was determined for every data set.
2) Link data to any publications based on or otherwise related to those data.
3) Make data more discoverable by allowing characteristics of data to be searchable across data sets.
FUTURE RESEARCH

1) Conduct similar studies with other neutron scientists to confirm results
2) Conduct studies of reuse “in real time”
ACKNOWLEDGMENTS

Thomas Proffen, Oak Ridge National Laboratory
Shawn Martin, Doctoral Student, Indiana University
UNDERSTANDING PERSPECTIVES ON REUSING NEUTRON DATA AT OAK RIDGE NATIONAL LABORATORY
DEVAN RAY DONALDSON, PH.D., AND SHAWN MARTIN

Email: drdonald@indiana.edu @hoosierdevan