The Status of Research Data Management in the United States

Challenges, Opportunities, and Emerging Trends

Dr. Martin Halbert, UNT Dean of Libraries

Open Data Day

University of North Texas, Denton, TX

Saturday, 21 February, 2015

UNT UNIVERSITY OF NORTH TEXAS

Overview

- Findings of DataRes Project analyzing changing national trends in research data management (RDM)
- New national developments and opportunities in RDM today

The Emerging Prominence of Research Data Management (RDM)

- 21st Century has been marked by a ***new sense of urgency around RDM*** topics such as managing, preserving, and providing access to data, especially in research communities
- Congress chartered the 2000-2010 ***National Digital Information Infrastructure and Preservation Program (NDIIPP)*** shepherded by the Library of Congress
- In 2003 ***NSF began issuing reports*** noting an urgent need to build national RDM capacity
- Many organizational recommendations and position papers, such as those of the ***Association of Research Libraries*** (2006) and the ***National Academy of Sciences*** (2009)
- ***Calls & mandates for better data management*** measures were issued by NIH, NSF, NEH, and the White House Office of Science and Technology Policy (OSTP)

UNT UNIVERSITY OF NORTH TEXAS

DataRes Project

- The 2011-2013 DataRes Project was funded by a 2011 grant of $226,786 from the Institute of Museum and Library Services 21st Century Librarian (21CL) program
- Goal: investigate the current status of research data management in universities and how the library and information science (LIS) profession can best respond to emerging needs.
- DataRes was a collaboration between the University of North Texas Libraries, the UNT College of Information, and the Council on Library and Information Resources.
- Paired with the iCAMP curriculum redesign project, another IMLS 21CL 2011 grant of $624,663 to UNT to assess educational needs and develop new shared curricula to train new LIS professionals seeking to fill data management positions

UNT UNIVERSITY OF NORTH TEXAS

Methodologies Used

DataRes Project

- Surveys of institutional data management policies, and views of individuals concerning research data management
- Textual analysis of agency requirements and institutional data management policies
- Focus groups of agency officials, university admins, and librarians
- Engaging community experts in producing a peer-reviewed Council on Library and Information Resources (CLIR) monograph discussing the future of research data management

iCAMP Project

- Competency analysis of advertised jobs in data management
- Curriculum redesign based on competency analysis

Textual and Focus Groups Analysis of Agency Requirements

Requirements analyzed

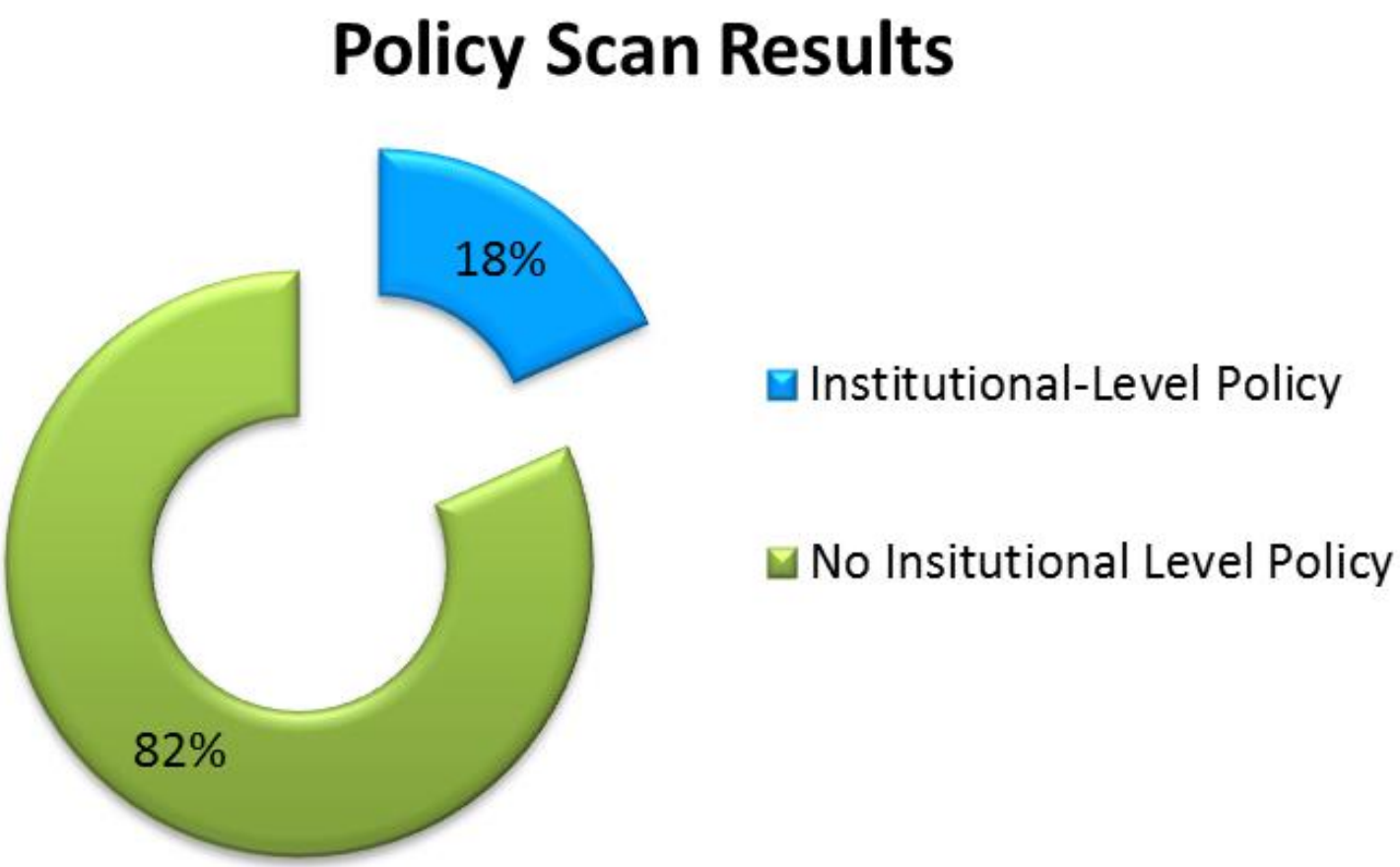
- National Institutes of Health "Final NIH Statement on Sharing Research Data."
- National Endowment for the Humanities, Office of Digital Humanities, Data Management Plans for NEH Office of Digital Humanities Proposals and Awards.
- National Science Foundation Award and Administration Guide. Chapter VI.D.4

Key findings

- There was significant variation across the different agencies in terms of emphasis and assertions
- Both disciplinary foci and historical level of emphasis on data sharing in the agencies was apparent
- Focus groups revealed a ***high degree of skepticism*** among many research communities that requirements would be enforced

UNT UNIVERSITY OF NORTH TEXAS

Notable Findings: Institutional Policy Scan



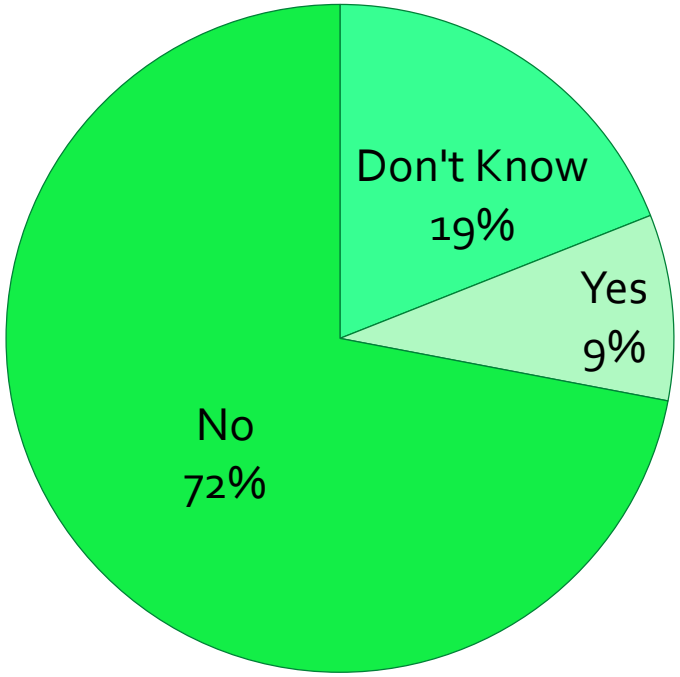
Text Analysis: Institutional Policy Scan

- 38 institutional policies identified and analyzed
- Policy language was often weakly assertive, example:

> *"The University recognizes the importance of data sharing in the advancement of knowledge and education."*
>
> (University of New Hampshire "UNH.VII.C.9")

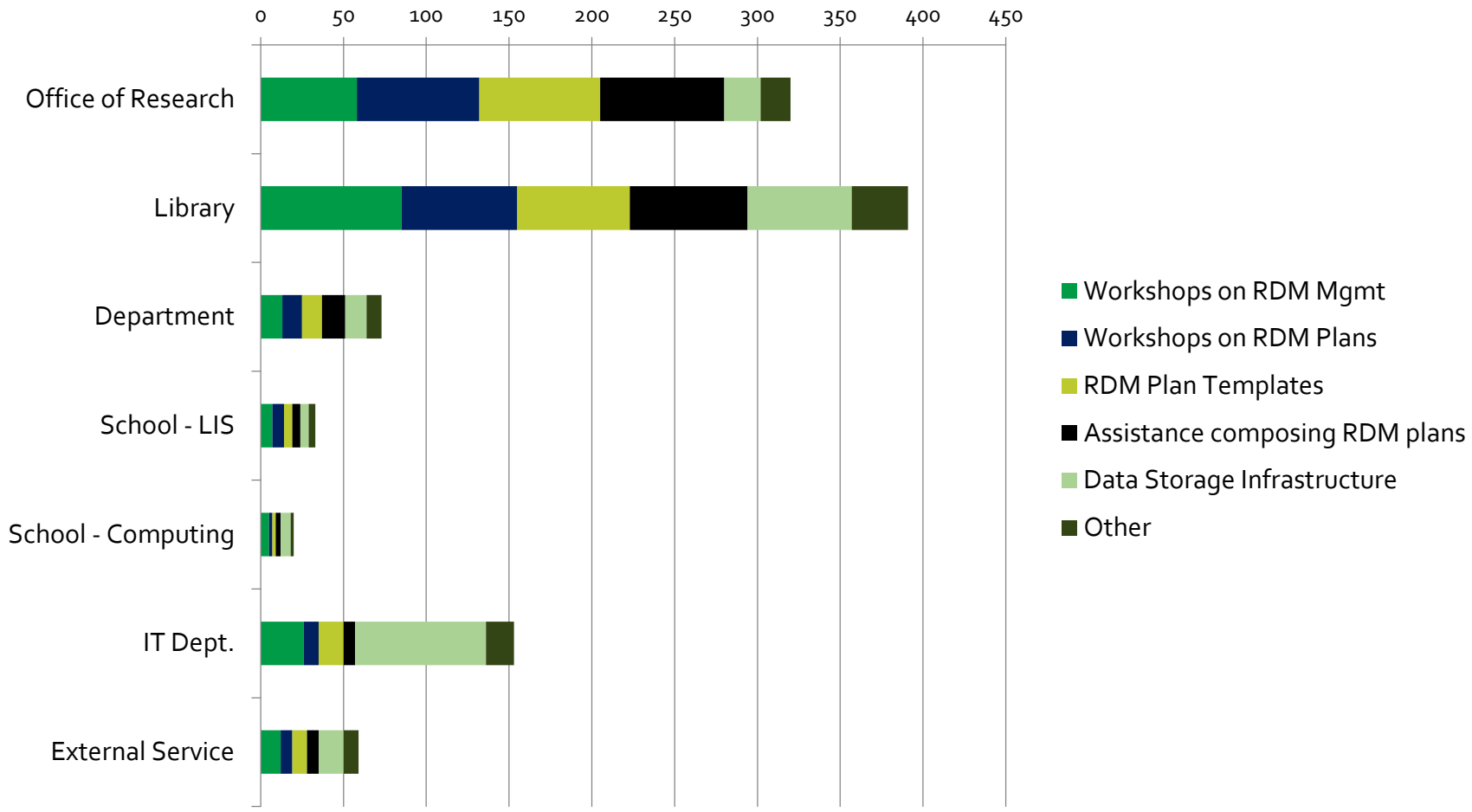
Notable Findings: Survey of Individuals

Does your institution have a policy governing the retention and sharing of research data? (231 responses)



UNT UNIVERSITY OF NORTH TEXAS

Institutional Support Chart



UNT UNIVERSITY OF NORTH TEXAS

The Denton Declaration: An Open Data Manifesto

- Developed by national gathering of university administrators, technologists, librarians, researchers, and other stakeholders gathered to discuss and articulate best practices and emerging trends in research data management
- Declaration is a statement of prescriptive assertions and values concerning research data management
- Bridges the converging interests of these stakeholders and promotes collaboration, transparency, and accountability across organizational and disciplinary boundaries

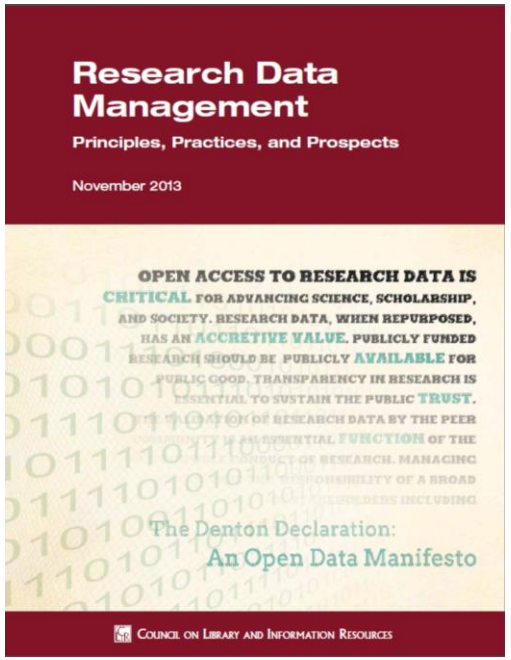
http://openaccess.unt.edu/denton_declaration

UNT UNIVERSITY OF NORTH TEXAS

DataRes Key Findings

- Major disconnect between assertions of the importance of research data management (by both agencies and individuals) and actual practice
- Much more prescriptive guidance and requirements will be necessary to actually encourage disciplines to take RDM requirements seriously in evaluating funding applications

http://www.clir.org/pubs/reports/pub160



UNT UNIVERSITY OF NORTH TEXAS

Key Findings (cont.)

- New competencies being requested for contemporary data management jobs are indeed significantly different from past
- While curricula can be redesigned, it is exceedingly difficult to find and recruit qualified instructors for updated curricula
- Institution-level policies are driven by practice, not the other way around
- Research Data Management is not a single university department issue, nor is it the purview of a single discipline
- Collaboration, domain knowledge, and infrastructure are all key to the success of any RDM response

Current Developments

Obama administration FY16 Budget lays out the administration's priorities, including (in pp. 74-75):

- "…invest in ***efforts to open up Government-generated assets, including data and the results of federally funded research and development***—such as intellectual property and scientific knowledge—to the public."
- "In demonstrating its commitment to open data, the Administration has developed performance metrics to ***measure agency progress in reaching open data goals***"
- "The Federal R&D enterprise must continue to support fundamental research that is motivated primarily by an interest in ***expanding the frontiers of human knowledge and diffusing this knowledge through open data and publications***."

(Perhaps as a result?) This month two agencies, the U.S. Agency for Healthcare Research and Quality (AHRQ) and NASA, followed the National Institutes of Health (NIH) to ***release their plans for ensuring public access to articles and data resulting from funded research***, as required by the February 2013 White House directive.

NIH Initiatives

- The National Institutes of Health (NIH) is in the process of standing up a contributions to open research data, notably their ***Big Data to Knowledge (BD2K) initiative***
- NIH is seeking to establish "a consortium of Commons compliant resources that will be used by investigators through science driven funding initiatives that seed the Commons."
- They are evaluating this pilot project to determine if sustainability, cost effectiveness and innovation can be achieved or anticipated.

Associate Director for Data Science (ADDS)

The Associate Director for Data Science (ADDS) leads the development of the overall NIH vision in Data Science and coordinates across the 27 Institutes and Centers in support of biomedical research as a digital enterprise. The mission of the ADDS office is to foster an ecosystem that enables biomedical research to be conducted as a digital enterprise that enhances health, lengthens life and reduces illness and disability.

Phil Bourne Blogs

RDA Fellowship

- The ***Research Data Alliance (RDA)*** invites applications for its ***newly redesigned fellowship program.*** The program's goal is to engage early career researchers in the US in Research Data Alliance (RDA), a dynamic and young global organization that seeks to eliminate the technical and social barriers to research data sharing.
- The successful Fellow will engage in the RDA through a 12-18 month project under the guidance of a mentor from the RDA community.
- The project is carried out within the context of an RDA Working Group (WG), Interest Group (IG), or Coordination Group (i.e., Technical Advisory Board), and is expected to have mutual benefit to both Fellow and the group's goals. Fellows receive a stipend and travel support and must be currently employed or appointed at a US institution.
- Applications due April 16, 2015 !