

The Problematic Future of Research Data Management: Challenges, Opportunities, and Emerging Patterns Identified by the *DataRes* Project

Dr. Martin Halbert, UNT Dean of Libraries
8th International Digital Curation Conference
Amsterdam, Netherlands
Wednesday, 16 January, 2013



DataRes Project

- The DataRes Project was funded by a 2011 grant of US\$ 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 of.
- DataRes is 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 US\$ 624,663 to UNT to assess educational needs and develop new shared curricula to train new LIS professionals seeking to fill data management positions

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

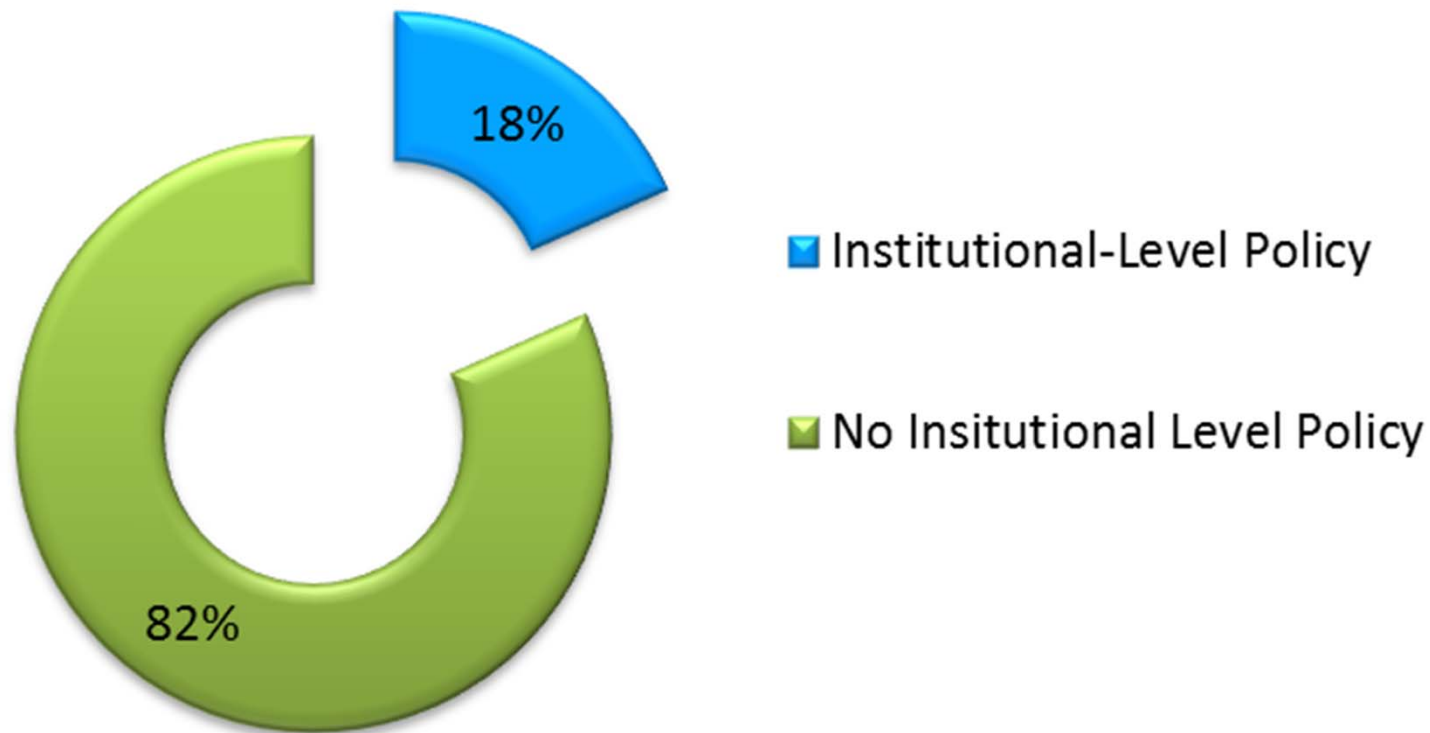
Key findings

- 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



Notable Findings: Institutional Policy Scan

Policy Scan Results



Text Analysis: Institutional Policy Scan

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

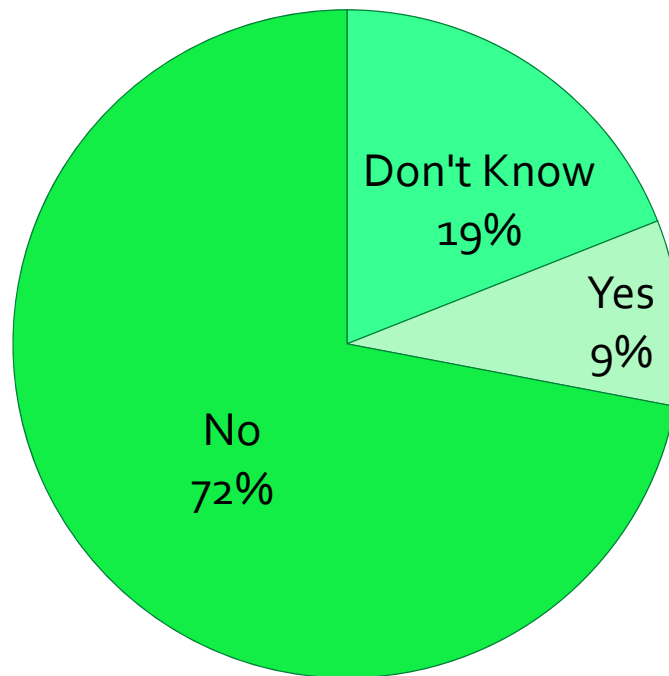
"The University recognizes the importance of data sharing in the advancement of knowledge and education."

(University of New Hampshire "UNH.VII.C.9")

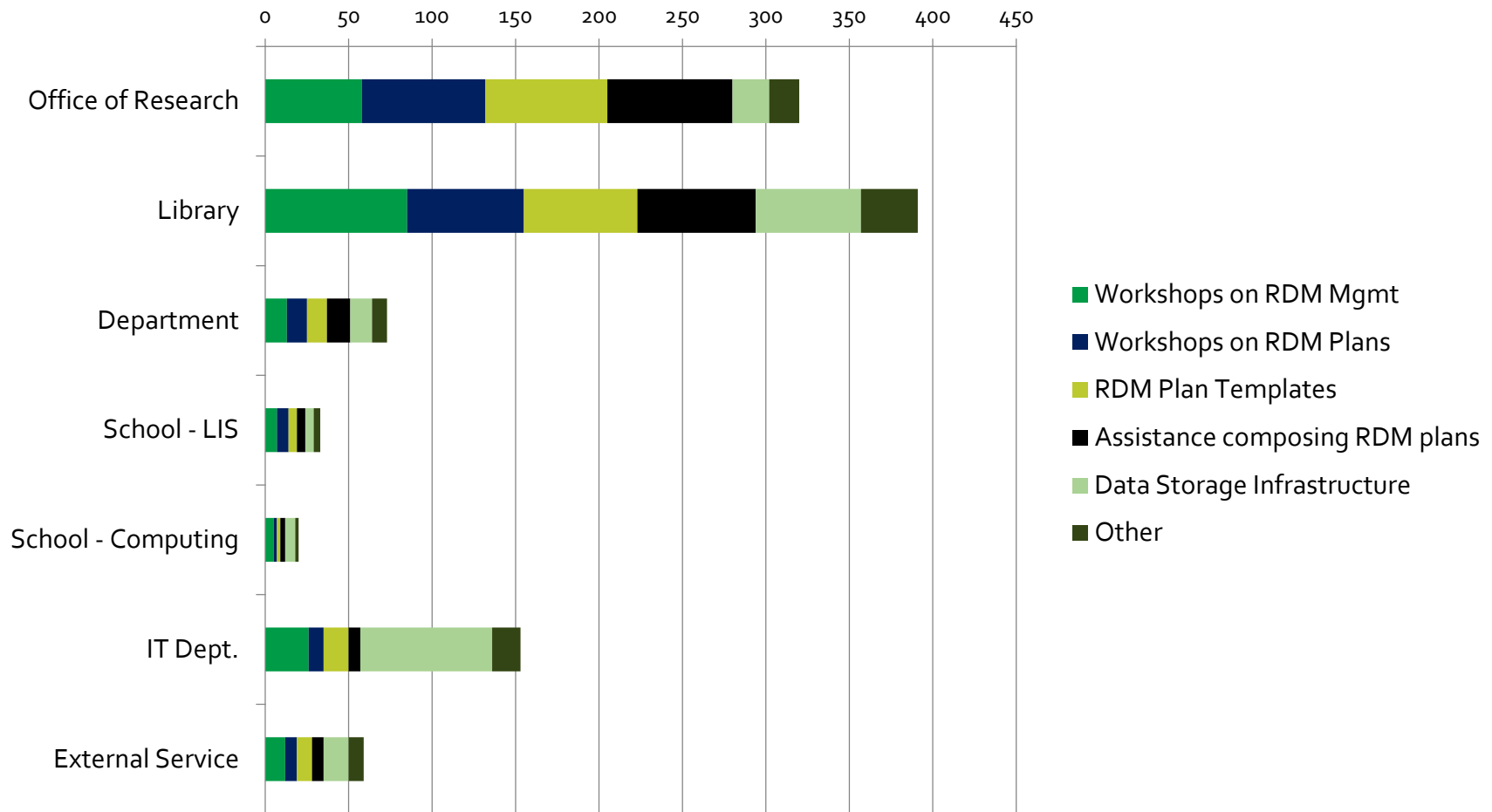
- An index of known policies, with links to the project data set will be available in February 2013 at:
<http://datamanagement.unt.edu/policies>

Notable Findings: Survey of Individuals

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



Institutional Support Chart



Secondary Survey to be Conducted in 2013

- Intended participants: VPs of Research, Deans, High-level Administrators
- We feel we got a very clear picture from the librarian perspective with the initial survey
- Questions were raised through the course of our survey analysis and research that we feel only institutional administrators will be able to answer

The Denton Declaration: An Open Data Manifesto

http://openaccess.unt.edu/denton_declaration

- 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

Synergistic iCAMP Project Findings

- 110 job advertisements were collected for analysis between October 2011 and March 2012
- Used to identify new competencies required for data management jobs
- New core curriculum of 4 courses designed and being taught for the first time in 2013
- Utilizing a “teaching library” model for practical training

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

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

Key Findings (cont.)

- Institution-level policies are driven by practice, not the other way around
- Research Data Management is not a single 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