SENATE COMMITTEE ON INFORMATION TECHNOLOGIES (SCIT)
MINUTES OF MEETING
October 7, 2014

[These minutes reflect discussion and debate at a meeting of a committee of the University of Minnesota Senate; none of the comments, conclusions or actions reported in these minutes represent the views of, nor are they binding on, the Senate, the Administration or the Board of Regents.]

[In these minutes: Charge review; Data Management Policy discussion; Public Access to U of M Data]

PRESENT: Jim MacDonald (chair), John Butler, Bradley Cohen, Michelle Driessen, Kristy Lashbaugh, Kate McCready, Karen Monson, Benton Schnabel, Yuk Sham, Zachary Shartiag, Tom Shield, Eric Watkins

GUESTS: Claudia Neuhauser, director, University of Minnesota Informatics Institute, Office of the Vice President for Research; Lisa Johnston, associate librarian - Public Access to U of M Research Data

OTHERS ATTENDING: Bernard Gulachek

REGRETS: Sean Conner, Tisha Turk, Diane Willow

1. WELCOME AND INTRODUCTIONS

Mr. MacDonald called the meeting to order and asked for introductions.

2. DATA MANAGEMENT POLICY

Claudia Neuhauser, director, University of Minnesota Informatics Institute, Office of the Vice President for Research, attended to consult with the committee on the Administrative Policy: Research Data Management: Archiving, Ownership, Retention, Security, Storage, and Transfer. Before the meeting she provided members a draft of the policy and the FAQ, and the Procedure: Transferring Research Data. She then provided members with background information on the policy:

- Development of the policy began in June 2013 when the University had an internal audit of OVPR and OIT with respect to research data storage. There were two particular recommendations made for the policy:
    - Create a process to determine the University's data storage needs
    - Clarify authority and responsibilities for data management
- May 2014 - A committee was created and had representatives from: OIT, University Libraries (1 from UMD because this will be a system-wide policy), College of Biological Sciences, College of Science and Engineering, School of Public Health, and institute

directors. The committee reviewed policies of other universities and consulted with others responsible for creating policies, for example:

- Provost Hanson
- Office for Information Technologies
- Office of the General Council
- Sponsored Projects Administration
- Clinical and Translational Science Institute (a major data consumer)
- Council of Research Associate Deans
- Administrators
- Faculty groups
- Senate committees: Faculty Consultative Committee, Faculty Affairs Committee, Senate Research Committee
- Feedback from all system campuses

She then walked members through the draft policy and highlighted the following points:

- The Responsibilities section contains clarification of the roles of the following areas and individuals:
    - VP for Research, VP and CIO, VP for Health Sciences, and the University Librarian/Director of Libraries
    - Academic Health Center (AHC)
    - Campus Library Directors (Excluding the Twin Cities – see University Librarian)
    - Colleges, Schools, and System Campuses
    - Office of the Vice President and Chief Information Officer (OVPCIO)
    - Office of the Vice President for Research (OVPR)
    - Principal Investigator (PI)
    - Sponsored Projects Administration (SPA)
    - UCCS Committee
    - Units other than Colleges, Schools, and System Campuses
    - University Librarian
- Definitions included in the policy:
    - Research data: Recorded factual material commonly accepted in the scientific or scholarly community as necessary to validate research findings, excluding preliminary analyses, drafts of scholarly or scientific work, plans for future research, peer reviews, communications with colleagues and physical objects (e.g., laboratory samples).

        - She explained that “physical objects” are done separately as research materials and are covered under different policies.
    - Principal Investigator (PI): The individual or individuals primarily responsible for and in charge of a research project.
- Ownership
    - University owns the data unless there is a prior agreement
    - Students own their data if it is part of their academic work and this aligns with copyright law. If they are employees, the University would own the data, this includes Ph.D. students because they are paid.
    - Fellows that are paid, but not by the University, are still treated as employees
    - Volunteers’ data is owned by the University
    - Co-ownership between a faculty member and student is possible, but there is a form that a student can sign signifying that they agree to have no rights to the data.
    - Ownership is impacted if substantial University resources are given to the student. A standard computer would not be considered a substantial resource.
- UCCS Committee will be created to discuss how the data should be stored.
    - College level solutions would need to be developed in many cases.
    - Reports back regularly to the VP for Research, the VP and CIO, the VP for Health Sciences, and the University Librarian/Director of Libraries (depending on campus), who, in turn, assign primary responsibilities for implementation to respective units, and will appoint the committee.
- What if a PI leaves? Or dies?
    - When a PI leaves, the University still owns the data or the University can relinquish ownership. The administering unit, preferably the dean, has to decide how the research is continued or stored. If the PI wants to take a copy, there has to be agreements made as to how the data is managed by the PI going forward. Sponsored data has agreements that would dictate this process.

Ms. Neuhauser emphasized that it is important for the PI to clarify ownership for those involved. Members then had the following discussion:

- Even students that are receiving credit, if they are being paid they are considered employees.
- Discussions should be done with students before to clarify these roles.
- Research data is place independent for example it can exist in Google Drive or a notebook.
    - Members discussed the issues related to Google account deactivation that results in the loss of data. There are administrative procedures for if a person leaves the University and their Google account is deactivated.

    - It is the responsibility of the PI to ensure the data can be accessed after they leave and also for backing up the data.
- The policy establishes that the University will need to be responsible and accountable for the data management past the funding of the grant. The policy does not need to determine how this will occur.
- Mr. Gulachek explained that policies have a Senate review schedule and the policy matrix can be found here: http://policy.umn.edu/prod/groups/president/@pub/@policy/@op/documents/policy/upolicy_senmatrix.pdf
    - Ms. Neuhauser added that since the policy draft has not been finalized it is not ready to be voted on, but comments could be collected for consideration.
    - Mr. Gulachek emphasized the importance of establishing this policy because once it has been finalized it is clear that the University has a responsibility in this space. Specific solutions are not provided because they risk becoming antiquated, which would make the policy outdated.

Closing the discussion Ms. Neuhauser asked members for their reactions to the policy via email. She explained that she would bring final comments to the Policy Advisory Committee. Members agreed that there is a need for the policy.

3. PUBLIC ACCESS TO U OF M RESEARCH DATA: launching new repository and curation services

Mr. MacDonald introduced speakers Lisa Johnston, associate librarian, and John Butler, Associate University Librarian for Data and Technology, University Libraries. Ms. Johnston is the lead of the Data Management and Curation Initiative (DMCI). They have collected data to inform planning and development of services to address the needs related to research data management. Mr. Butler explained that though the administrative policy on research data management is not finalized, they have been developing the services to support researchers for many years. He added that the research data policy is a positive step toward recognizing the need for the clarity of roles in providing services to support researchers.

Ms. Johnston then provided members an update on the Libraries development of services for data management as well as repository services to make research data created here accessible. She used a PowerPoint to highlight the following points:

- Researchers at the U of M Need Services for Sharing and Preserving Research Data
    - NSF and NIH have had policies that ask researchers to provide a plan for how they will share their data publicly and make it available for reuse and analysis. Moving forward nearly all federal funders will require similar plans.

    - Conducted a pilot of data curation services. They sought to answer questions such as: What is involved in making the data ready for public access? How can data be prepared for people to understand it?
    - They successfully made data available through the Digital Conservancy.
- User Needs Assessment on Campus
    - Partnered with CLA, surveyed faculty in fall 2012, and asked the needs of researchers in research and data management. The same survey was also sent to AHC and CSE in the summer and fall of 2014.
    - They obtained research data management plans that were attached to grants submitted by UMN faculty and accepted by the NSF since 2011.
- CLA Survey Results
    - Preserving and sharing data were identified as areas where services are needed.
- AHC Survey Results
    - Services for preparing a data management plan was expressed as an area where more services are desired.
- DMP Review Project
    - Received 183 out of 400 plans sent to NSF since 2011.
    - Preparing private data is an issue, but many believed that it does need to be shared.
- We Already Offer Data Management Support
    - They offer training, DMP consultation, assistance to graduate students, archiving and preservation.
- Data Curation Facilities Reuse
    - The Libraries is launching a new service that will involve the review of data curation to help researchers make the data reusable, resulting in a greater impact.
    - Data curation steps may include appraisal, ingest, arrangement and description, metadata creation, format transformation, dissemination and access, archiving, and preservation of digital research data.
- Libraries' Data Management/ Curation Initiative
    - Objective 1: Offer data curation services that are appropriate for UMN research data.
    - Objective 2: Partner with other Campus Units so that we can scale.
- Gaps in UMN Research Data Lifecycle will address issues in the following areas:
    - DMP consultation
    - Metadata Consultation
    - Data Management best practices training
    - Data curation and repository services
- Date Repository of the University of Minnesota (DRUM)
    - Utilize existing repository technologies for cost savings/efficiencies (DSpace, open source software)

    - Custom upload form and metadata schema for research data
    - Curation workflow allows for review of data before openly available
- DRUM Benefits to Researchers
    - Free and Open Access - flexible access options
    - Persistent Access - Persistent links and identifiers (DOIs) make it easy for others to cite your data.
    - Increase visibility - Analytics to track how often your data is viewed and downloaded
    - Meet grant requirements - Comply with federal mandates for data management planning and sharing
    - Long-Term Preservation
- Timeline to Launch
    - DONE Service model articulation and 89+ functional requirements. (Feb-May 2014)
    - In Progress Repository development, testing and branding (May-Sep 2014)
    - Upcoming User testing and implement communications plan (Oct 2014)
    - Launch of DMCI Services ~ Tentatively October 20, 2014
- Partner with Campus Units
    - Create a referral network to help create the service around the user
    - Redesign web site for users and better display the relationships with other units on campus
    - There is an iCOP for data management and she encouraged members to join if they are interested

Mr. Butler explained that the Libraries recognize that they cannot proceed alone. The following are large-scale collaborations that involve massive digital storage and preservation management:

- Massively-scaled Collaborations
    - HathiTrust Digital Library (b.2008)
    - Portico (b. 2002)
    - DPN - Digital Preservation Network (2012)

Mr. Butler then opened the discussion for comments and questions:

- Where is the data stored? Storage can often limit access to data, depending on the method and capacity.
    - University Archives has been invested in the digitization of objects like lab and field notebooks.
    - There are instances when data exceeds what the University can provide. There are some review processes dictated by the draft policy that would require review at the VP level and this could result in greater investment.

    - Transferring data from Minnesota Supercomputing Institute (MSI) to the Digital Conservancy. Sometimes it does not make sense to move the data, however MSI will close accounts once the project is complete.
- Mr. Cohen explained that the policy draft is an indication that the University is showing leadership and attempting to create solutions at the institutional level.

Hearing no further business, Ms. MacDonald adjourned the meeting.

Jeannine Rich
University Senate Office