

Research Policy Committee Minutes

Friday, November 30, 2018

2:30-4pm

**Topic: The Role of the Library in Research, Teaching and Scholarship**

Fleming 4025

Chaired by Francine Dolins

**Members present:** Francine Dolins, Mimi Dalaly, Sami Malek, Irene St. Charles, Jessie Lee, Jairam Menon, Adam VanDeusen, Nick Harris, Tim Guetterman, Ashley Kalinski, Francois Beaufay, William Close, Kate Eaton, William Schwartz, Jake Carlson, Marisa Conte, Yi-Su Chen

**Absent:** Albert Liang, Niccolo Biltramo, Austin Glass, Nocona Sanders

- I. Minutes from September 2018 meeting – approved with minor changes (added the names that were omitted and fixed the mis-spellings).
- II. **Guest Speakers:** **Marisa Conte:** Assistant Director for Research and Informatics  
**Jake Carlson:** Director of Deep Blue and Research Data Services  
**Topic: The Role of the Library in Research, Teaching and Scholarship**
- III. **Marisa:** [Assistant Director for Research and Informatics, Medicine, Taubman Library]
  - a. The ways in which the library supports research is varied and multi-level. Marisa is involved in Translational Science. She works towards various goals. One is to find good, reliable information to further research and patient care. She also helps researchers to gain access to information and databases, to help organize their own, and to disseminate the information more widely.
  - b. The data management plan in health sciences is based on the current NIH data sharing and management policy. Marisa helps researchers write data management plans and helps them gain access to repository facilities that are necessary or helpful for NIH grants. She assists in improving work flow on grants specifically.
  - c. NIH provisions and requirements for data sharing – Marisa helps researchers with best practices in data management, sharing and to demonstrate the impact of the project. Also, in re-structuring data to share and decide what to share. The objective is to help distinguish and consolidate what can be shared without burdening the project and what can be shared to impact the field. Marisa also assists in how to gain credit for data sharing (ie., reputational enhancement).
  - d. Marisa has also contributed to a ‘Policy Piece’ on the National Library of Medicine, focusing on what comprises data sharing and the policies of journals, which are to make data available to others to reach broader audiences for additional analyses and further output from those data sets.

- e. Marisa also has a counterpart in the Medical Library whose focus is on Clinical aspects of the health sciences.

IV. **Jake:** [Director Deep Blue and Research Data Services]

- a. Jake explained the importance of data sharing and how the UM Library, particularly, Deep Blue, can assist researchers with this.
- b. Jake explained how data are shared so that it is available and not lost, so that it enhances a researcher's reputation and has impact in the field.
- c. In his role, Jake helps to provide structures for sharing, and longevity for impact. This creates a connection to others and information flow to others – for information sharing
- d. There is the issue of metadata regarding copyright, for which Jake can assist researchers to protect their datasets.
- e. Jake is also available for data lifecycle consulting; he can assist to let researchers' know what else is available to be able to use data sets that already accessible. He also can assist to show how data sets should evolve and grow, help create data management plans, and provide documentation for how the data sets develop and grow. This provides a focus on sharing data and the longevity of the data sets.
- f. The Library's repository for published papers, books, etc is **Deep Blue**, for UMichigan authors.

**Deep Blue Data** is a repository of data sets within a protected environment that is Sustainable and managed.

For **Deep Blue Data** – it allows for metadata analyses; it can permit others to use these data sets to cite, and therefore treats data sharing as normal part of the scientific process of scholarship and is recognized and acknowledged for career and promotion purposes. Metadata it has as much functional quality for understanding and use by others.

- g. There is a strict curation review for any data going into Deep Blue Data for preservation. Not all data sets are accepted, and some are thinned out after a number of years based on set criteria. Deep Blue will keep dataset for 10 years – Deep Blue reviewed and then if decide to remove, will contact researcher about what they want to do.
- h. There are also connections between publications and the Deep Blue Data set so that they can be viewed as parts of the same project. These data can be harvested by Google and aligned for metadata searches.
- i. There are three levels of preservation:
  - 1. Open format
  - 2. Proprietary – no underlying schematic – can preserve what was given but can't do more.
  - 3. Had been proprietary but now is open.

- j. Size of data sets uploaded into Deep Blue Data: There are limitations for file sizes in storing? For really large sets – 2 gigabytes or less, these can be uploaded by the researcher themselves.  
If sets are more than 2 gigabytes, then they need to be uploaded by the Librarians for storage, but only sets up to just under a terabyte. For scans over a terabyte – the question is how to store these large files/data sets (such as large sets of MRI scans used in a research project, which take up a large amount of bytes). This would need a plan because there is no present solution for storage of such large amounts (over a terabyte).
  - k. Data Refuge from government agencies: the Lumos group has been salvaging data.
  - l. Databrary through New York University has extensive capabilities for data storage, sharing, paper writing, etc.
  - m. R programming allows for data to be analyzed and updated automatically in a paper while being worked on by multiple authors.
  - n. A suggestion of using *Software Carpentry* to learn how to use R programming. There are workshops at the UM to help researchers to use this.
- V. Announcements: Kate Eaton: Three points from the Research Administration Advisory Council (RAAC) meeting the day before:
- 1. - ORSP developing a portal to help with writing a grant.
  - 2. - An update on MIDAR (dashboard for faculty to get all info on projects and grants in one place).
  - 3. – Due dates for proposal review in ORSP begins in April 2019. There are three levels of review possible, full, partial and just submission on time.
- VI. The meeting was adjourned at 4pm.