DWG Activity Report
2010 - 2011

Authored by the Data Working Group
Rebecca Reznik-Zellen (chair)
Jessica Adamick
MJ Canavan
Maxine Schmidt
Steve McGinty
Contents

Completed Tasks ........................................................................................................................................... 3

Faculty Interviews ...................................................................................................................................... 3

Graduate Student Focus Group .................................................................................................................... 3

DMP Template ........................................................................................................................................ 3

Web Audit .............................................................................................................................................. 4

Communication ...................................................................................................................................... 4

Consultation ........................................................................................................................................ 4

Current Efforts ......................................................................................................................................... 5

Workshops and Education for Fall 2011 ................................................................................................... 5

Feedback from NSF PIs ............................................................................................................................ 5

Article in Journal of e-Science Librarianship ............................................................................................. 5

Future Projects ......................................................................................................................................... 5

Evaluation of infrastructure for management and curation of data sets ................................................... 5

Microservices for data curation ................................................................................................................ 5

Recommendation of cross-campus committee for data ........................................................................... 5

Acknowledgements .................................................................................................................................. 5
The Data Working Group began its work in March of 2010, just prior to the first NSF announcement of its new requirement for data management plans. This report summarizes the activities of the Data Working Group from its establishment in early 2010 through June of 2011.

The Data Working Group has lost important and active members over the course of the last year: Kelcy Shepherd, Aaron Rubinstein and Bob Stevens. We hope to begin repopulating the group soon with people who are interested in data management and/or represent important stakeholder groups, such as the incoming Digital Repository Resident.

**Completed Tasks**

**Faculty Interviews**
During August and September 2010, the Data Working Group invited a handful of faculty from various disciplines to participate in short, casual interviews regarding their data management practices. Seven of thirteen invited faculty participated in the interviews. They came from Computer Science, Polymer Science, Natural Resources/Public Policy, Civil and Environmental Engineering, Psychology, and Geology. In general, each faculty member has a very unique story to tell about their research data and how they collect and store it. At the same time, they collectively demonstrate some of the larger issues surrounding data management as a whole. The overall picture of current data management on campus is that it ends at publication, and faculty data management strategies vary widely up until the point of publication. See [Data Working Group November 2010 Report](http://www.library.umass.edu/services/services-for-faculty/data-management/data-management-plan-guidance/).

**Graduate Student Focus Group**
During Open Access Week 2010 (October), the Data Working Group co-hosted “I Can Has Data?” a focus group for graduate students with the Graduate Office. Seventeen students participated, from Mechanical Engineering, History, Education, Polymer Science, Chemical Engineering, and Sociology. As with the faculty, graduate student data management practices vary across disciplines. These students are primarily responsible for the collection, documentation, and management of data for their research projects, although they report receiving no formal training on this topic. Their primary concerns for data management centered on data storage and backup, versioning, and security. See [Data Working Group November 2010 Report](http://www.library.umass.edu/services/services-for-faculty/data-management/data-management-plan-guidance/).

**DMP Template**
In Spring of 2011, the Data Working Group created a Data Management Plan template and posted it to the Libraries’ Web site. The document is a general template designed to inform the development of a data management plan. Its elements respond to the requirements of the National Science Foundation mandate for Data Management Plans and are based on several similar guidance documents. See [http://www.library.umass.edu/services/services-for-faculty/data-management/data-management-plan-guidance/](http://www.library.umass.edu/services/services-for-faculty/data-management/data-management-plan-guidance/).
Web Audit
In the Spring of 2011, the Data Working group performed a web audit of peer and model institutions to survey library approaches to data at other universities. The web audit evaluated the following areas of library data work: infrastructure, services, organization, and marketing (see Table 1). A list of peer institutions used for evaluation at a university-wide level included Indiana University Bloomington, Iowa State University, Rutgers University, SUNY Stony Brook, University of California Santa Barbara, University of Colorado Boulder, University of Connecticut, University of Delaware, University of Maryland, and University of Nebraska Lincoln. Model institutions were identified based on the Data Working Group's research on library data initiatives: Purdue University, Massachusetts Institute of Technology, University of Minnesota, Johns Hopkins University, Cornell University, Oregon State University, University of Virginia, and the University of Wisconsin-Madison. Based on the public information available through peer and model institutions’ web sites, three levels of research data support services have been identified: education, consultation, and infrastructure. See Data Management Web Audit Report Executive Summary.

Communication
The Data Working Group engaged in several on campus and external communications activities on the topic of data management.

- Prepared a letter to the campus regarding the NSF mandate signed by Jay Schafer and Mike Malone
- Created and published data management Web pages on the Libraries’ Web site under Services for Faculty
- Initiated two meetings with Bruce McCandless of the Office of Research & Engagement
- Participated in Working Group Confab, January 2011
- Hosted meetings with Mt. Holyoke and UConn to discuss data management on our respective campuses, and visited UMass Medical for same
- Presented on the NSF mandate and data management to the Research Council, Research Library Council, FSUCECC, and All-Staff
- Presented a poster on our approaches to data management at the 3rd Annual University of Massachusetts and New England Region e-Science Symposium

Consultation
We have consulted with the following faculty on data management and/or data management plans.

- Brian Lickel, Psychology: in person consultation on data management plan
- Dave Hart, Computer Science: email consultation on data management plan
- Steven Schneider, Astronomy: email request for DMP template
- Elisabeth Selkirk, Linguistics: in person meeting with ScholarWorks on data management
- Sue Roberts (via Mary Fechner), Chemistry: email consultation on data management plan
Current Efforts

Workshops and Education for Fall 2011
Currently, the Data Working Group is planning a series of workshops for Fall 2011 that are targeted to graduate students and to faculty members. We also plan to meet with Liaison librarians to present the materials that we have available, in the event that their faculty express an interest in data management or data management plans. Additionally, we are also planning a panel on Data for the 2011 Open Access Week activities.

Feedback from NSF PIs
At this time, the Data Working Group has circulated an email to those faculty who have submitted a Data Management Plan to the NSF (list courtesy of Bruce McCandless) as a follow up measure and effort to initiate conversations about the challenges that they have encountered in creating their data management plans.

Article in Journal of e-Science Librarianship
The Data Working Group was invited to submit an article for the inaugural issue of the Journal of e-Science Librarianship based on the poster that we presented at the 3rd Annual University of Massachusetts and New England Region e-Science Symposium. This issue will be published in September of 2011.

Future Projects

Evaluation of infrastructure for management and curation of data sets
In the coming year we plan to begin the evaluation of different platforms or services that are available for data management, including cloud services.

Microservices for data curation
We plan to investigate data curation microservices, as presented by Joan Starr of the California Digital Library (http://www.slideshare.net/joanstarr/toward-making-data-management-easy).

Recommendation of cross-campus committee for data
Several times in our discussions over the last year the idea for an intra-institutional body to discuss and drive activities around data management has come up. In the coming year we will push this idea forward.

Acknowledgements
As chair, I would like to thank all the members of the Data Working Group, past and present, for their hard work. They are: Jessica Adamick, MJ Canavan, Steve McGinty, Aaron Rubinstein, Maxine Schmidt, Kelcy Shepherd, and Bob Stevens.