

Data Management/Curation Task Force, Quarterly Report (June 2013)

Members

Hannah Norton (Co-Chair), Laurie Taylor (Co-Chair), Rolando Garcia-Milian, Denise Bennett, Val Minson, Joe Aufmuth, David Schwieder, Blake Landor, Erik Deumens, Robert Ferl (replacing Stephanie Gray), and Cecilia Botero (ex-officio)

Background

The Data Management/Curation Task Force was charged in late 2012 to assess needs, make recommendations, and develop support for the role of the Libraries in campus-wide data management and curation. The charge includes specific advisory and operational activities (full charge: <http://www.uflib.ufl.edu/datamgmt/index/TFCharge.aspx>).

Overview of Current Activities

The group meets every two weeks on Wednesdays at 1pm, rotating the meeting locations across the Health Science Center Library, Library West, and the Marston Science Library.

For the second quarter (April-June), the group has focused on sharing information and on building the timeline and plan of activities for future quarters.

Work to date includes:

- Sharing information, building communication and collaboration among Task Force members, and with external groups and constituents
- Reviewing best practices, models, and resources from peer institutions
- Facilitating focus groups and discussions on data management including: CLAS IT (4/26); English and Film Studies (4/29); SPOHP (5/1, 6/4, and 6/18); African Studies archives (6/17); Poli Sci (6/18); Chemistry (7/10); and upcoming with Soc and Psych (7/16)
- Planning and delivering training and training materials for librarians, faculty, and students regarding data management, research, and reference, including these workshops: Zotero (5/14); Intro to HPC (713); and others: <http://www.uflib.ufl.edu/datamgmt/index/WorkshopsTrainings.aspx>
- Testing the Dataverse Network with the Southeastern Universities Research Association (SURA); see DVN report for more information
- Completing the initial customization of the DMP Tool; attending webinars for capacity for enhanced support; liaised to support the libraries in completing the MOU for use
- Creating and distributing promotional materials and information related to data management, including identifying existing events and activities as related to data management and promoting them as such: <http://www.uflib.ufl.edu/datamgmt/index/DMCSpeaker.aspx>
- Creating additional resources, linked from the Task Force website

Findings/Recommendations from Quarter 2

During this second phase of work by the Task Force, it was clear once again that campus data management and curation needs are extensive and critical as more people continue to seek out and contact the Task Force for assistance and guidance, and by the high level of interest from all. This quarter, we focused on the focus and facilitated group discussions, finalizing the survey with IRB approval for release in the next week, testing the Dataverse network and identifying a more extensive testing scenario (see additional report), and building collaboration among many groups

Future Quarters

Work is expected to continue at a rapid pace for the next quarters.

Specific upcoming activities by time period include:

July-August:

- Participate in the ASERL Data event on August 6 (Val, Rolando, and Laurie) to build community, reporting back to the group
- Participate in the UF Big Data even on August 7
- Submit a funding proposal to ORCID to incorporate ORCID identifiers for SobekCM and for usability support and training for new interface to easily allow UF authors to create ORCID ID from within the IR, dLOC, and other hosted digital libraries. (PI: Val Davis; Co-PI: Mark Sullivan; Co-PI: Laurie Taylor; Project Team: many)
- Collaborate to ensure information on data and contacts are included in new trainings:
 - Graduate Students, Aug. 13-16 (<http://graduateschool.ufl.edu/admission/orientation>)
 - Faculty, Aug. 14-15
 - Post docs, Sept. 24
- Presenting within the libraries on activities to date and expected future activities
- Begin to implement more extensive Dataverse network test with IFAS-NFREC and Chemistry
- Collaborating to have data questions included in the responsible conduct in research survey
- Planning training and other supports, based on information gathered during Focus Groups and other activities
- Discussing, planning, and implementing workshops and types of workshops to support UF librarians and other members of the UF community who may soon provide data services
- Training for the fall – initial brainstorming
 - We would like to plan a 1-2 day workshop for late September, targeting graduate students. Late September is expected to be a good time because new students will be somewhat settled in by then.
 - *Potential topics:*
 - Introductory panel (members of the task force) talking about overall context, DMPTool, finding data, metadata, storage, the IR, etc.
 - Introduction to HPC (Matt Gitzendanner)
 - Breakouts on particular elements related to data analysis (individuals listed in parentheses are ideas about who could be involved or volunteers from the Task Force):

SPSS (Dave Schweider), SAS (Lou Ann Cooper), GIS (Joe Aufmuth), Google Refine (Val Minson), using Excel for research data (Pandora Cowart), copyright/permissions/data sharing (Christine Fruin), data visualization (Joseph Murphy, Chad Juehring, grad students at Digital Worlds)

- Perspective from/profile of a researcher (Erik Deumens or Rob Ferl)
- Possible workshop: Writing a data management plan using the DMPTool and example DMPs from the IR@UF with Michelle Leonard (est. for 2-3 hours)
- Possible workshop: Primer on Data Management
 - Possible draft outline for 2 hour version:
 - 1 hour on overview
 - What is data management? Why does it matter to me?
 - Walkthrough on process steps and concerns
 - Designing from the start (data management begins before research)
 - Creating data
 - Refining data
 - Publishing/archiving data
 - Citing data in CV and from others
 - Sharing data (done with proper publishing and DCC lifecycle)
 - 1 hour working through a specific example for field/research type
 - Expanded primer within 2 day workshop, co-taught with teaching faculty in-field
 - Expanded primer within lab-style courses as with research and methods courses, etc.
- Drafting proposed plans for future activities, to be submitted in next Quarterly Report and for proposed presentation within the libraries on activities to date and expected future activities.
- Continued re-drafting the task force's website (<http://www.uflib.ufl.edu/datamgmt>) to eventually serve as the umbrella public site for data management planning.
- Developing a template for creating IR records of datasets for discovery.

Fall 2013:

- Implement more extensive Dataverse network test
- Implementing training program
- Draft proposed recommendations as whitepapers for review/approval/implementation to include: Recommendations for the Libraries' campus-level role in support of data management and curation; proposing a corresponding framework and resources for library support of the data life cycle; recommending the role of the institutional repository and research computing in storing, finding, and accessing working and final data, and linking publications to supporting data; and, recommending a framework for liaisons and subject specialists to incorporate data instruction and consultation into their workflows. Possibly include:
 - Recommend that all libraries add a tab to existing LibGuides for relevant data repositories to find and deposit data

- Recommend and promote best practices for submitting posters to the IR to support related data and as part of data management on its own.
- Quarterly presentations /information distribution within the libraries on activities

Resources

- DMCTF site: <http://www.uflib.ufl.edu/datamgmt>
- Related wiki: <http://uflibdatacuration.pbworks.com/>
- Email list: datamgmt-l@lists.ufl.edu