Minutes from University Library Committee 4/1/2014  
Approved April 29, 2014

Present: Steve Barkin (Law), Allison Kaplan (SLIS), Laura Rudquist (Graduate Student), Dan Klingenberg (Chemical and Biological Engineering), Baha Balantekin (Physics), Ed Van Gemert (Ex-Officio), Jody Hoesley (Librarian, non-voting member), Karl Broman (Biostatistics & Medical Infomatics), Ron Harris (English), Karen Britland (English), Nilay Vaish (student representative) Eileen Cullen (Entomology), Larry Nesper (Anthropology), Mary Trotter (Theatre & Drama) Alan Wolf (CIO’s Office), and Michael Enyart (Librarian, non-voting member)  
Also present: Julie Schneider (Ebling Library), Leslie Mayo (GLS), Doug Way (GLS), David Null (Archives), Nancy Graff Schultz (GLS), Lisa Wettleson (Steenbock).

K. Britland opened the meeting and turned it over to E. Van Gemert for his report. Prior to his report he introduced Lesley Moyo, Associate University Librarian for Public Service (Services and Space within the GLS) has been on campus since March 1st. E. Van Gemert also introduced Doug Way, the new Associate University Librarian for Collections and Research Services. Doug Way’s portfolio will include special collections, allocation of funds, selection, and coordination of those activities. The ULC has spent a good amount of time talking about the management and quality of the GLS collections. The University Libraries has to consider their collection strategy going forward and the University Library Committee will be of great help in forming that strategy. In addition to these Associate University Librarians, the other Associate University Librarians are L. Konrad, Associate University Librarian for Technology Strategies and Data Services, the fourth is Nancy Graff Schultz Associate University Librarian for Administration. E. Van Gemert mentioned that with the addition of Lesly Moyo and Doug Way, this completes his leadership team.

At the February ULC meeting Heather Weltin presented the print management strategy that the GLS will use going forward. This strategy utilizes the electronic resources that we own in perpetuity and reliance on shared print agreements (CIC, CRL, etc.). To date there have been two sessions where this strategy has been presented. The first one was intended for campus library staff. It went well; most staff understands the electronic nature of many library resources in this day and age. The meeting also highlighted the different needs of the sciences versus the arts and humanities. The second session for faculty and staff and it was well attended. K. Britland also attended this session. The audience was primarily faculty from the Humanities. The faculty at that meeting wanted to make sure that the library understood the disciplinary needs in the humanities for print material. E. Van Gemert is certain that the GLS understands the importance of print and he is also confident that things are changing. Going forward E. Van Gemert wants to adopt the goal to be able to retain a high level of services. Part of that high level of services is a guarantee that libraries will get researchers what they need on this campus in a timely manner. The difference is that (especially in terms of serials) that while UW Madison Libraries will have collection ownership, we will not have it all the time on our shelves as it has been in the past. E. Van Gemert felt that it was a good discussion. He felt that there was skepticism among the arts and humanities faculty that the print management plan would serve them in good stead, but they understand they there are decisions that have to be made regarding our print collections. K. Britland commented that faculty wanted to let the
library know that how some scholars use print does not always lends itself to digitization. Such things as paratext and marginalia are rarely captured in digitization projects. This whole area of scholarship could compromised by some of the decisions that the library makes. There may be different strategies that are tied to different disciplines. How the sciences do something may not be applicable to the arts and humanities. E. Van Gemert mentioned that the print management plan that is going forward is focused on serials not monographs. E. Van Gemert mentioned that it is important that we maintain dialogs with students and faculty regarding our service model and print management strategies.

At the next meeting which will be the last of the academic year, E. Van Gemert would like to present his view on where he thinks the GLS is headed and get feedback from the committee. He would also like to summarize what has accomplished during the current academic year and some of the strategic areas that the GLS is moving into and will need the support and consultations of the University Library Committee. The GLS will need to request dollars from the University Administration; it will be different request than it has been in the past. It has to be targeted. In discussions with the University there is some indication that they are willing to help. However, we need to use all of our allies when we request additional library funds in the next fiscal year.

E. Van Gemert also reported on the search for a new provost. E. Van Gemert was part of the Provost’s Executive group that got to meet with each candidate separately in addition to their public presentation. He mentioned that it was easy to see why these candidates were chosen to be interviewed out of the larger pool. E. Van Gemert feels that the Chancellor may be making the decision in the near future. E. Van Gemert will be a producing a report for the current provost but it most likely be used by the incoming provost. In that report he will plant the seed regarding the funding needs of campus libraries.

K. Britland asked for any corrections or additions in the minutes. Eileen Cullen mentioned that she attended the last meeting but her name is not on the minutes as being in attendance. M. Enyart will make the necessary changes. K. Britland asked for the minutes to be approved with the necessary changes. Committee approved the minutes.

K. Britland then turned the meeting over to Julie Schneider and Alan Wolf to lead a discussion of research data management services that are currently being provided on campus. Deb Helman, Director of the Wendt Library Commons, was scheduled to be part of this presentation but is out ill today; J. Schneider is Director the Ebling Library and Alan Wolf, Assistant CIO for Advanced Computing Infrastructure. Here is the link to their slides.

While Alan and Julie were setting up, E. Van Gemert wanted to frame the discussion. Everyone is cognizant of the emergence of the need of information management of research data. From the Library’s perspective we need to know how to coordinate a service package, what are we doing in this space, and why it is important. We are looking for feedback from the University Library Committee in this area.

Alan Wolf started out the presentation by talking about research data management, how it is changing on campus and what role the library has in this arena. Definition of research data: The information generated/acquired by campus researchers
Changing nature of data collection:
Volume - Most of campus researchers were never trained to deal with the volume of data that they are now able to collect.
Format – Transitions to digital data storage have new opportunities and risks that our researchers were not prepared for.

DATA LIFE CYCLE

Project Planning (Pre-Award)
- Proposal support
- Service selection
- Data discovery
- Proofs of concept
- Data disposal
- Data management plans

Project Implementation (Active Award)
- Storage
- Data management tools
- Data workflow management
- Metadata support
- Data sharing and discovery

Project Conclusion (Post-Award)
- Long-term storage/archiving
- Transformation and preservation
- Data sharing and discovery
- Data disposal

Alan shared a campus example.
Kevin Eliceri, of the Laboratory for Optical and Computational Instrumentation (LOCI), provides services to people who use optical microscopes. His lab produces thousands of images a year. Since he started on this campus he has had a goal of sharing all of these images with the practitioners and researchers in this field. In his facility he has worked with SLIS interns to manage all of this image data. He recently started working with Peter Gorman of the DCC to make these image data available and discoverable in the campus catalog. This is a good first step. Kevin will continue to work on such problems as data migration from one format to another.

Baha Balantekin (Physics) asked A. Wolf what he meant by data, since in his view it is a very generic term. B. Balantekin went further and said that in Physics many of their research projects generate gigabytes of data by the second. A. Wolf responded by saying a good data work flow model is that one employed by the CERN Collider. It produces Terabytes of data for each run. B. Balantekin stated that the data can only be shared by collaborators and those collaborators decides what becomes public. A. Wolf said the data needs to be sorted (some data
is deleted if there is no event) and stored. Eventually the data will be shared and some of will be public upon the publication of studies. All of this is part of a data management plan. B. Balantekin asked if by data did A. Wolf mean publicly available data. A. Wolf stated that we need to support the data needs of researchers independent of whether the data is public or not. Some of it may not be public. The flu researcher on campus is a prime example. B. Balantekin responded that it is research services, not necessarily campus libraries. A. Wolf said that this is one of the questions that need to be answered. What is the Library’s role in data management?

J. Schneider began her part of the presentation talking about all of the activity in the last 6 years at the National level, in terms of public policy and legislation. The first piece of legislation was in 2007 with the NIH public access policy which was publications only. The proposed legislation over the last three years has included the requirement about sharing data, but none of this has been passed into law as yet. Some of the proposed legislation will expand the NIH policy to all federal departments that receive over 100,000,000 in extramural funding. The policy that demands a data management plan originates with the Executive Branch through the Office of Science & Technology Policy. The directive requires all federal departments with over 100,000,000 in extramural grants to provide public access to publication and data that were the results of federal funded research. How this will implemented is still not clear. Took out last sentence

Repositories that go hand in hand with the data storage and management plans are also instrumental. Some of those are institutional (e.g. Minds at UW) and other might be provided by funding agencies or the disciplines or associations of those disciplines.

Open Access or Public Access has increase over the last few years. There are currently 46 institutions that have open access requirements.

Data Citation: There are international associations that are working on the process for data citations and what format. Guidelines on data citations will be forthcoming. There was a question on what does this mean. J. Schneider said that the idea is that one can cite a data set like one cites a journal article.

Researcher ID: There is a growing need for such an ID. A number of funding agencies require a research id in their applications. Given there are a number of researchers with the same name, a researcher ID will be of great help. Sort of like a DOI for individual researchers.

Current Campus Activities:

Data governance (Administrative/Educational): Various departments (DOIT, CIO’s office) are talking about this. They are looking at legislative requirements. What is the best way to manage that data governance? A. Wolf said the discussions of the data is educational not research. Campus is thinking about a data governance policy. Who maintains the data and how that is accomplished, who has access, record retention etc. Will research data be exempt from those policies?
Role of the University with compliance:
University compliance with the legislative and OGSP policies regarding publications is very high. When the new NIH publication policy was formulated the libraries and SLIS ramped up quickly to provide support to researchers so that they would be compliant in this new requirement. Although it is not a rule as yet, many are anticipating that as publications have become publicly available, the data that is collected by extramural grants will also have to be made available publicly. In this arena campus libraries, CIO office, DOIT, and SLIS developed a research data service as a means to provide assistance to researchers on campus. At this point research data services have provided expertise in data management planning, data storage, and data citations. All of these member groups are building services under the aegis of research data services but is in need of information as to which services will be needed the most.

E. Van Gemert asked the following question:
What is the compliance rate on this campus for NIH publications? J. Schneider stated that the current compliance rate is 90%. E. Van Gemert then asked if the inclusion of data management plans as part of the funding application have had a detrimental effect on the campus compliance rate. Again, J. Schneider responded that the data management plan is not part of the NIH application process but it is part of the NSF application process. She went on to say that there is no way of knowing if a NSF application was denied funding due to a poor data management plan.

E. Van Gemert mentioned J. Schneider and others librarians have been instrumental in assisting campus researchers to be compliant with the NEH policy on publications. The data part of research is altogether a different story. For one thing the amount of data is huge, scalability is a large issue.

E. Van Gemert asked how receptive the researchers on this campus are to the efforts of the research data services that include DOIT and libraries. A. Wolf feels that the researchers that know about the service, the feedback has been very positive. However, data management for NIH and NSF applications is still a relatively new phenomenon.

E. Van Gemert asked if it is fair to say that libraries we are not staffed to handle the data management need. J. Schneider responded that in terms of publications we have a number of people in place that can handle questions and research administrators that are also very knowledgeable. In terms of data management we are not yet at the same level. A. Balantekin (Physics) mentioned that this is the first time that he has heard that this campus provides some data management services. A. Wolf commented that this is a common problem with researchers not always being aware of the services that are available to them on campus.

L. Wettleson asked is the Department of Agriculture part of this scenario. Both Julie and Alan did not know but thought it probably was governed by the same policies.

K. Britland asked about security. Is there part of data management plans that deals with security? Is there a nationwide conversation that focuses on security of data? A. Wolf felt that this might be more in the area of cyber security. It is one of the things that campus will have to examine. The data management tools are just a template on how to manage their data. Some of the other tools actually contain data and protection of that data is important (e.g. flu
research). For some researchers using some of these tools may need additional security. Risk assessment needs to be done for each project.

Tools:
DMPTool: Data management plan. Allows you to plug in information and it provides a data management plan.
ELN (electronic lab notebook)
   - Campus RFP – intent to award stage
   - Cloud solution: compliant with campus policies and key requirements
Digital Libraries / Archives / Repositories: Minds@UW
Campus IT Initiatives
   - Campus Data Centers
   - ACI – Advanced Computing Initiative
   - Shared Preservation Service – link to Digital Preservation Network
   - Possibilities for greater efficiency, compliance, security, etc.

Developing a research data management strategy:
   - Having such a strategy is more important than just compliance
   - How we can manage the scale of data that is generated, how to we make it secure, how do we share that data.
   - Having a management plan allows UW to increase its efficiency and competitiveness.

Question: Where is the Library’s role in Data Management? The good news is that the library is not alone in handling this issue.
Partners
Current
   - UW-Madison Libraries
   - DOIT
   - Office of the CIO
   - Graduate School – set policies
   - School of Library & Information Studies – train next generation of information managers
   - Faculty & Researchers – early adopters: Telling us what they do want and do not want.
Future
   - Advanced Computing Infrastructure
   - Research Deans
   - Research & Sponsored Programs
   - CIC – how do we leverage with the CIC for data storage, data tools
   - Other groups looking at campus data management issues
   - Other?
Here is a list of questions that J. Schneider and A. Wolf have to assist in moving the data research services forward.

- How could this impact your work?
- What kinds of services would be helpful?
- What would motivate the use of these services?
- What types of campus policies would need to be addressed?
- How do we market/brand this “data” service that welcomes all campus researchers to use them?

Dan Klingenberg (Chemical and Biological Engineering) talked about a data management issue that arose from a faculty meeting. The result is that the department decided to institute their own mirrored drives behind their firewall. This is a very distributed model. He asked Alan what are the problems with that type of service. A. Wolf commented that the question just resolved about storage. A data management plan encompasses more than just storage. Here are some of the other issues:
  - Performance of storage
  - What kind of systems that would lie on this storage that helps meta-data to help with retrieval
  - How to you share that data (difficult behind firewalls)
  - Customization possible
  - Security/Protection/Back-up

D. Null of University Archives stated that one of the issues of data management on this campus is enforcement. There are a number of policies that determine how long a document or data should be maintained. In a distributed environment enforcement of these rules become very difficult and as a result increase campus legal liability.

E. Van Gemert wanted to explore A. Wolf’s statement that data management is not just about science. Can Alan give an example? A. Wolf responded that there are some massive digitalization projects (Haithatrust). To Alan it is data not text. There are many digitization projects in the Humanities where a good deal of print is converted to digital for research and analysis. Large image collections are another example.

E. Van Gemert stated that the data management services group needs a larger audience. There needs to be a broader campus discussion. It is importation to be careful about the library’s role in data management. A. Wolf feels that librarians have the training and background that would lend itself to forming the campus strategy on data management.

Karl Broman (Biostatistics & Medical Infomatics) worries about the lone computational researchers in a non-computational department. There will be more computational researchers in the humanities. Data Research Services will be an important resource for those researchers. The main role of the library is that there are a large number of aspects of data management and distribution where librarians have the key understanding and skills (metadata, discovery, etc.). It is Important that the library is a part of this, even if libraries do not have a leading role.
K. Britland mentioned that the next ULC meeting will be at the end of April. Members that want items on the next meeting should email her. She thanked J. Schneider and A. Wolf for their presentation.

Meeting was adjourned

Respectively Submitted,

Michael Enyart