2014

Report of the ALCTS Scholarly Communications Interest Group Meeting, American Library Association Midwinter Meeting, Philadelphia, January 2014

Doug Way
Grand Valley State University, doug.way@gmail.com

Report of the ALCTS Scholarly Communications Interest Group Meeting, American Library Association Midwinter Meeting, Philadelphia, January 2014

The ALCTS Scholarly Communications Interest Group's meeting at the 2014 ALA Midwinter meeting featured presentations on research data programs at two universities. Mary Bergstrom presented a case study that examined the evolution of the research data curation program at the University of California-San Diego, while Michele Claibourn and Ivey Glendon described research data services at the University of Virginia.

In the first presentation, "Pilots to Program: UC San Diego Research Data Curation Pilots and the Library Research Data Curation Program," Bergstrom, presenting on behalf of her colleagues Matt Crichlow, Arwen Hutt, Declan Fleming, David Minor, and Don Sutton, outlined the evolution of UC San Diego's data curation program from a pilot project to a full-fledged program.   In 2008 the university began a campus-wide needs assessment that included an examination of the needs of researchers in terms of cyberinfrastructure. This needs assessment was followed in 2009 and 2010 by reports that provided rationales and business plans for operationalizing a campus-wide research cyberinfrastructure. From the start the Library was included in these conversations and reports, and in January, 2011, when a two-year pilot to develop a Research Cyberinfrastructure (RCI) launched the Library was included in that program.

The RCI pilot aimed to use existing infrastructure whenever possible, including storage at the San Diego Supercomputer Center, the Chronopolis digital preservation network, and the Digital Asset Management System (DAMS) at the UC San Diego Library. The pilot also took advantage of existing tools and services that were also in place at that time, such as the Library's metadata consultation services and workshops, and the EZID service from the University of California Curation Center for assigning Digital Object Identifiers (DOIs).

In her presentation Bergstrom described the partners and projects that the RCI pilot worked on over those two years. Projects included collaborations with The Brain Observatory, the NSF OpenTopography Facility, the Levantine Archaeology Laboratory, the Scripps Institution of Oceanography, and the Laboratory for Computational Astrophysics. During the pilot the RCI team met regularly with these pilot partners, often subdividing into smaller groups based on the needs and interests of the partner group. Bergstrom described the work done with each partner group, such as the development of metadata and workflows for data transfer, which she said was an iterative process that often entailed each group receiving personalized attention. Bergstrom said that in some cases this close collaboration was not seen as a sustainable model in the long-term because of the high level of assistance provided, but was sustained over the course of the pilot so the RCI team could learn as much as possible. The pilot provided a number of take-aways. For example, the team learned that the Library's existing DAMS was not adequate to handle the data that was coming in. As a result the DAMS were reconfigured, allowing them to work with more complex data sets.

The research data programs the Library offers today following the conclusion of the pilot include metadata services for complex research data, repository services

for data using the Library DAMS, long-term preservation of data in the Chronopolis digital preservation network, DOIs for data, and training on data management. The UC San Diego Library made changes to its organizational structure during the same time the pilot project was occurring, with research data services moving from Technology Services to Collections Services, with programs like metadata services and preservation services. The Library is also increasing its staffing on the research data curation program.

Looking toward the future the Library is looking to develop the consistent use of a researcher profile tool across campus. Bergstrom also discussed interest in electronic lab notebooks, improving data ingest functionality, implementing visualization and integration tools, data management tools, establishing data information literacy standards, working with social media data, and then generally focusing on communication, education, collaboration and assessment.

Questions following the presentation focused on how the program was being assessed and whether quantitative or qualitative methods were being used, as well as on outreach related to research data services. In her responses Bergstrom discussed how DOIs could be used for assessment, and how liaison librarians are assisting with marketing and outreach, as well as challenges that still exist in making their services known.

In the second presentation of the session, "Data Services as Information Services: or, Old Wine, New Bottle," Claibourn and Glendon discussed the Research Data Services (RDS) program at the University of Virginia. Claibourn began the presentation by discussing the changing research data environment, the potential for the use of research data, and the challenges that exist today. She went on to explain how the research data services program at the University of Virginia Library is designed to respond to these challenges and take advantage of the expertise that already exists in the Library in areas such as reference consultation, the organization of knowledge, and the preservation and curation of information. Its objective is to help researchers manage, preserve, find, access, and use data, seeing data services as a kind of information services and data literacy as a kind of information literacy.

In making the case for why a library should be the hub of research data services at a university, Claibourn argued that as the use of data has grown across disciplines and has also become more interdisciplinary, there is a growing need for one place to provide centralized services. At the University of Virginia the library had existing expertise and programs related to data services, such as data management planning services and metadata expertise. This made it easy to fill in the gaps in existing expertise to further support and expand these programs. Once these pieces were in place the Library successfully made the case to university administration that it should be the face of data services on campus.

Claibourn described how RDS aims to provide researchers help in a variety of ways and at different stages of a project. This includes helping people at the start of a project with planning, not only when they are writing grants, but also in designing data-collection efforts and statistical analyses. RDS also provides support in the collection of data, both with identifying and applying appropriate metadata, but also with the acquisition of needed resources. RDS assists with the use of data, such as

helping researchers with the use of GIS tools or statistical software. Finally, they provide services that assist with preserving and sharing data, helping users think about how to make their data reproducible and reusable.

RDS has been able to do these things by developing a structure that integrates knowledge and services from multiple units within the Library. Claibourn said that while the Library has an RDS department, they did not want the singular responsibility for RDS to live in that one department, but instead wanted RDS to be integrated throughout the Library. To this end they have a steering committee that pulls together people from across the Library to provide direction, facilitate ongoing communication, and to promote referrals and cooperation. All of this serves to increase collaboration, take advantage of existing expertise, and to highlight the connections between the work being done in the different departments and units.

Glendon shared some specific examples of how the University of Virginia is implementing these services. She said these services fall into three main categories: tools, collections and resources, and teaching and consultation. In the area of tools, the focus has been on assisting users data management planning tools, statistical software packages, and GIS data tools. The areas of collections and resources focus primarily on providing researchers with access to data collections. The Library has hired a data acquisitions librarian and is focusing on purchasing and providing access to data sets. Finally, teaching and consultation aims to provide generalized instruction and assistance. This has entailed maintaining a robust schedule of workshops, particularly with a metadata focus, but also working with researchers one-on-one.

To date the Library is pleased with the growth of these services. The Library held a variety of workshops in the fall semester with topics ranging from data management planning to metadata to the use of software and specific tools. The Library also saw an increase in one-on-one consultations with researchers. It held a data management bootcamp that was sponsored with four other Virginia universities, and has also begun hosting open office hours to promote the service and to reduce barriers to access.

In concluding their presentation, Claibourn discussed the RDS goals for the coming year. These include continuing to build the unit's identity and raising awareness outside the Library. The program has developed a newsletter and will be launching a new web page. They are also looking to expand and raise awareness by creating two physical research data "hub" locations where people can come when they are seeking help. They are working to expand participation within the Library, identifying those who want to develop more data-related skill sets, as well as hiring more people to support data services.   Finally, they are looking to develop additional services and partnerships, ranging from the development of secure data rooms to collaborations on big data projects.

Following the presentation, Claibourn and Glendon answered a number of questions regarding the availability of online materials related to their program, staff involvement in the program, content covered in their workshops, and whether the University of Virginia has a data repository. Claibourn responded that the Library does have materials online, but they are not centralized, which pointed to

the need for the centralized website she discussed in her presentation. Claibourn also shared that approximately 15-20 people were involved in the RDS, with 12 people being very actively involved in the program. In terms of interest, she said that there has been a mixed reaction, with some people very excited and becoming engaged and others worrying about the imposition of additional responsibilities. In response to the question about the content of workshops, Glendon said the focus has been to provide generalized metadata instruction and principles that researchers can apply to their own projects. Finally, Claibourn explained that the University of Virginia does not currently have a data repository, so at the moment they point users to subject repositories or tools like Dataverse.

Both of these presentations provided attendees with an excellent overview of two approaches to research data services. The standing-room only crowd at this session speaks to the growing interest in research data services, and likely to the pressures libraries are under to provide support to users in these areas. As libraries look to develop programs to meet local needs, the research data programs outlined in these two case studies can serve as solid starting points. Slides from both presentations can be found on the ALAConnect Website (http://connect.ala.org/node/66015).

Doug Way
Grand Valley State University
Allendale, MI