

EDITOR'S SUMMARY

The University of Colorado-Boulder expanded its digital services through the 2016 opening of its Center for Research Data and Digital Scholarship (CRDDS), a collaboration between the libraries and research computing. The aim is to provide a full range of data services, from analytics and visualization through curation, storage and preservation, in addition to education and consultation for both university and community members. Increased data discovery, reuse, access and publication are expected to be among the several advantages to this collaborative effort. The CRDDS will focus on promoting research data management mandated by federal funding agencies and journal publishers, cyberinfrastructure to gain full benefit from big data, education and training through courses and online modules, and digital scholarship to explore tools and methods used throughout the life cycle of a digital project. CRDDS plans to create a seminar and consulting space within the library in addition to digital outreach.

KEYWORDS

data curation
colleges and universities
research libraries
outreach services
scholarly publishing

The Center for Research Data and Digital Scholarship at the University of Colorado-Boulder

by Shelley L. Knuth, Andrew Johnson, Thea Lindquist, Debra Weiss, Deborah Hamrick, Thomas Hauser and Leslie Reynolds

In October 2016 the Center for Research Data and Digital Scholarship (CRDDS) was created with the mission of supporting the University of Colorado Boulder (CU Boulder) and the general community in developing an advanced data infrastructure and of assisting CU researchers with developing data skills and advancing their digital scholarship. CRDDS is an active collaboration between the CU Boulder libraries and research computing, the group that provides large-scale computational and data infrastructure on campus. The center's mission includes data education, data quality consulting and assistance with data analytics, visualization, management, storage and preservation. CRDDS will build on existing services and skills across the CU Boulder campus to build a research and scholarly data ecosystem that is cross-disciplinary and connected to many local, regional and national collaborations. The main purpose of the center is to assist researchers with the following tasks:

- How to determine what data exists and how to retrieve it

- How to prepare data for use
- How to analyze and visualize data using the latest tools
- How to preserve and manage data for proper future use and reuse.

It is anticipated that, by providing these services to researchers, the center will advance the state of the art in digital data and scholarship, by specifically contributing to

- discovery and reuse
- access and publication
- management, curation, and preservation
- analysis and visualization
- training and education.

CRDDS will be a partner in research and scholarship projects primarily for campus groups, but also for the community by providing support services, new training and education approaches, and consulting. Initially the center is organized around four initiatives, each with an initiative director: research data management, cyberinfrastructure, education and training, and digital scholarship. Each initiative is described below.

The Four CRDDS Initiatives

Research Data Management. The Research data management initiative of CRDDS builds off the existing research data services (RDS) group at CU Boulder [1]. This group, the first formal collaboration between the libraries

The authors are all staff of the University of Colorado, Boulder. Shelley Knuth and Thomas Hauser are associated with the Office of Information Technology while Andrew Johnson, Thea Lindquist, Debra Weiss, Deborah Hamrick and Leslie Reynold are from the library. They can all be reached at `firstname.lastname<at>Colorado.edu` except Andrew Johnson, whose email address is [Andrew.M.Johnson<at>Colorado.edu](mailto:Andrew.M.Johnson@Colorado.edu).

and research computing at CU Boulder, provides support for research data management on campus (<https://data.colorado.edu>). RDS provides support for compliance with data management planning and data sharing requirements for federal funding agencies and journal publishers. RDS offers one-on-one consultations, seminars and workshops to assist researchers, with about 40 consultations and 10 workshops per year [1]. RDS also administers the DMPTool (<http://dmptool.org/>) [2], creating 240 data management plans as of September 2016.

Cyberinfrastructure. The cyberinfrastructure initiative of CRDDS leads development and deployment of the central CU Boulder data infrastructure. This infrastructure includes storage, big-data analytics approaches and data management and curation software and applications. The PetaLibrary, the largest storage infrastructure at CU Boulder, is managed by research computing and is, in conjunction with the CU Boulder ScienceDMZ, the core of the current research data infrastructure of CRDDS [1]. CRDDS-affiliated research staff will work with researchers to provide consultation surrounding this and other storage services across campus. Support will also be offered for data management, analytics, visualization, cleaning and processing. The cyberinfrastructure initiative will also assist users in discovering knowledge that exists in big data and in leveraging cloud or other computing services for their data analysis needs.

Education and Training. This initiative will work closely with the other initiatives of CRDDS, as well as other campus groups, to ensure the CU Boulder campus training needs are met. This initiative will develop new or incorporate existing courses and informal seminars centered around data science and management skills. This

training program will also develop short online modules or videos about data science subjects designed to reach an even larger audience. In addition to the informal training approaches, several formal for-credit classes related to data are being developed to create a data-related certificate program. The tuition revenue from the certificate is anticipated to help fund the center.

Digital Scholarship. The digital scholarship initiative promotes exploration and integration of digital scholarship tools and methods – such as social network analysis, geospatial analysis, text and data mining, and digital exhibits – into research and teaching. This initiative will facilitate support networks for all stages of the digital project life cycle, from project planning to dissemination of scholarly outputs. It also will promote researcher engagement with cultural heritage data, especially digital primary-source collections, and offer guidance in scholarly communication, including authors' rights, open access and publishing in the institutional or other repositories.

Structure and Inclusion of Other Groups on Campus

The governance structure of CRDDS includes four initiative directors, two executive directors, an executive board, an advisory board and a set of affiliates. The initiative directors include representatives from both the libraries and research computing, while the executive directors are the senior associate dean of the libraries and the director of research computing. The executive board consists of the associate vice chancellor for IT, the dean of libraries, two deans from CU Boulder colleges or schools, and the vice chancellor for research.

CRDDS will strive to create an inclusive and cross-disciplinary culture to encourage participation from other groups across campus who are participating in data-related center activities. This effort will be particularly relevant through the advisory board and the affiliates program. The advisory board will consist of up to 12 members including a broad representation of stakeholders from faculty, students, staff and community advocates. The board will include at least five tenure-track CU faculty, one post-doctorate, one graduate student and one undergraduate student. Affiliates will include personnel across campus and within the community who are interested in formal and informal partnerships with CRDDS.

Future Work

As of this writing, the CRDDS has been an official entity for less than one month. The initial plans are to develop a consulting and seminar place within the main library on campus, with the intention of having this space available in early 2017. The seminar space will have a direct network connection into the ScienceDMZ to enable visualization of large datasets. Seminars will be broadcast to the larger community via streaming. Other plans are to develop the advisory board and set of affiliates. Plans to incorporate other campus groups are well underway. The center will play an integral role in the development of the PetaLibrary 2.0 and other data related campus infrastructure. ■

Resources Mentioned in the Article

- [1] Knuth, S.L., Johnson, A.M., & Hauser, T. (2015). Research data services at the University of Colorado Boulder. *Bulletin of the Association for Information Science and Technology*, 41(6), 35-38.
- [2] Strasser, C., & Cruse, P. (2013). The DMPTool and DataUp: Helping researchers manage, archive, and share their data. Unpublished position paper for the Research Data Management Implementations Workshop, March 2013, in Arlington, Virginia.