# Barriers to Collaboration: Lessons Learned from the Data Curation Network

**Lisa R. Johnston**, Research Data Management/Curation Lead and Co-Director of the University Digital Conservancy, University of Minnesota Libraries, and Principal Investigator of the Data Curation Network Project

There are many barriers that prevent us from actively and equitably collaborating in meaningful ways. When we launched the cross-institutional Data Curation Network (DCN) project,[1] our team took conscious steps toward seeking out those barriers and working to find ways to overcome them. I will present those barriers here and note some ways that we are attempting to overcome our obstacles.

First, a bit of background on our project. Our vision for the Data Curation Network is to ensure that researchers, when faced with a growing number of requirements to ethically share their research data, are preparing and archiving their data in ways that make it findable, accessible, interoperable, and reusable (FAIR). Data curation activities—such as quality assurance, metadata/documentation creation, code review, and file transformations—support FAIR data publishing and sharing activities. But data curation can be costly, requiring advanced curation practices, specific technical competencies, and relevant subject expertise. For multidisciplinary institutions and nonprofit data repositories, the sheer range of data curation expertise required to perform these services well is an enormous challenge. The DCN takes a collective approach to data curation. By sharing our expert data curation staff across DCN partner institutions, we enable ourselves to collectively, and more effectively, curate a wider variety of data types (for example, discipline, file format,

> The Data Curation Network project brings together the perspectives of research data librarians, academic library administrators, and domain subject experts from academic libraries and general-purpose or disciplinary data repositories.

etc.) beyond what any single institution might offer alone.

The Data Curation Network project brings together the perspectives of research data librarians, academic library administrators, and domain subject experts from academic libraries and general-purpose or disciplinary data repositories. Our project began in 2016 with six partners and funding from the Alfred P. Sloan Foundation, and has since grown to include eight partner institutions including the University of Minnesota (lead), Cornell University, Dryad Data Repository, Duke University, Johns Hopkins University, Penn State University, the University of Illinois, and the University of Michigan.

Curation staff are the "human layer" in the repository technology stack who bring the knowledge and software expertise necessary for reviewing incoming submissions to ensure that the data stand up to the test of time and are optimized for reuse. We do this several ways. First, the DCN creates a platform for partner institutions to share our curation staff using a coordinated workflow that connects data sets to the appropriate expert for that particular data type (for example, GIS data, 3-D images, simulation data, etc.). Second, the DCN provides a community for professional data curators. By sharing tools, providing a pipeline for training data curators, and promoting data curation practices across the profession, the Data Curation Network aims to enrich capacities for data curation writ large. Third, the goal for the DCN will be to offer sustainable services and access to data curation expertise to end-users (researchers, libraries, journals, etc.) when none exist locally, for rare or infrequent data types, or in times of staff transition.

> " Curation staff are the 'human layer' in the repository technology stack who bring the knowledge and software expertise necessary for reviewing incoming submissions to ensure that the data stand up to the test of time and are optimized for reuse. "

To confront the challenges of collaboration, at the onset of our project we identified some specific barriers that might keep us from moving together toward a shared vision. We revisit these barriers annually and consider ways to reduce or eliminate these barriers. Some of the challenges that our project has faced include the following:

- **Institutional priorities and culture.** Each institution has different goals and priorities for how they approach data services. Institutional competition and internal competition (for example, tech transfer office goals at odds with library repository mission) could prevent DCN collaboration. Multi-institutional collaborations must deal with different institutional and local cultures.

  - Site visits are planned at each member institution to discuss the project goals and outcomes with institution administration.

- **Unvoiced concerns.** Are we doing a good job at onboarding new DCN members? Are we building curator buy-in? Or creating opportunities to voice dissenting opinions?

  - Regular in-person meetings have been one way to bring everyone in the DCN together. At these events we encourage multiple communication methods (for example, writing anonymous feedback and leaving it on the "ideas" table).

- **Indeterminable or unknown value proposition.** There is scant market research or literature to show that curated data are more valuable to researchers. What if our efforts are not valued or not well communicated? What if the costs outweigh the value? Demand for data curation is low, but metrics fail to tell the whole story.

  - Our research agenda includes white papers describing the value of data curation to funders and stakeholders and documenting the cost savings of collaborative data curation.

- **Complex and evolving ecosystem.** Data sharing requirements and norms are in flux. Data curation is only part of a larger

conversation about data sharing. Norms and best practices of curation are still forming.

- An early effort in our project was to research and document a shared glossary of data curation terms.[2]

- **Challenges of practical network design.** There is a tendency to over-engineer and create complex workflows. On the other hand, not everyone will "see themselves" in a more general workflow. There is a need to find balance.

  - Developing a framework for shared work is changing. Our goals have been to not change how local institutions do data curation, but to keep the DCN workflow modular and allow institutions to decide locally how to best incorporate a shared staffing network. There have been many trial-and-error opportunities.

- **Antiquated and limited view of libraries.** Libraries face skepticism about having a role in data services at all. Some curators don't want to "criticize" researchers' data.

  - Our planning phase spent a considerable amount of time holding focus group interviews with researchers to understand what data curation activities they find important and where our project could make the most impact.[3]

- **No sustainable funding model.** It is challenging to find and secure sustainable funding in an age of austerity. Within the cacophony of data projects and "membership fatigue"—being heard is hard.

  - In our current phase we aim to engage a sustainability consultant to help navigate these issues.

- **Easier to do it yourself.** Library work is often built around relationships. If we rely on others to perform complex data consultations with local researchers, what opportunities for strengthening relationships are lost? It may be better, easier (or perceived as such) to do all curation work locally.

  - A strong lesson learned from this project has been to keep

local control over how and when to engage support from the network.

- All communication to local researchers will be mediated by a local curator so that connections can be strengthened and maintained.

- **Transparency.** How do we communicate what we are doing with our local campus? What if we don't tell researchers that external staff are curating data? Do we lose trust? (Do they know when using 24x7 chat services that these are done by libraries in different time zones?)

  - This is an area we will be closely watching and assessing in our implementation phase over the coming year.

- **Unbalanced workloads.** Collaboration can mean more work for overburdened staff. How much local time/effort can be devoted to working on "someone else's" data? Participating institutions are at different places in their curation services (and expertise); at home institutions, what happens when one partner overuses shared resources?

  - A grant project model will protect us in some ways (for example, staff have a dedicated amount of time to spend on the project and we will use project management software to help us keep track). But maybe we will need to let go, be more flexible.
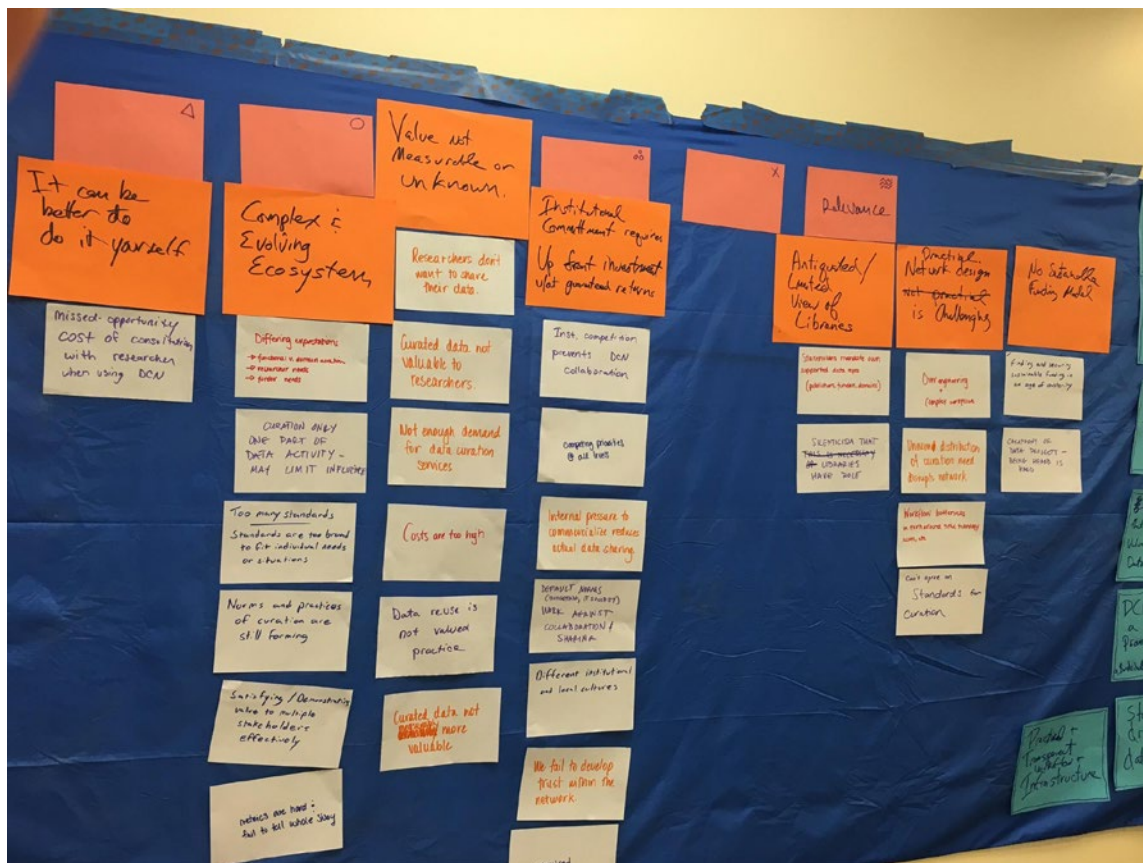
At the June 2016 kickoff meeting of the Data Curation Network, the team was led by a team collaboration expert to envision the goals for the project and to acknowledge the potential barriers. Notes from this meeting are published online.[4]

## Endnotes

1. Data Curation Network website, accessed December 5, 2018, http://datacurationnetwork.org.

2. Lisa R. Johnston, Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, and Claire Stewart, *Definitions of Data Curation Activities used by the Data Curation Network*, October 2016, retrieved from University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/188638.

3. Lisa R. Johnston, Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, and Claire Stewart, "How Important is Data Curation? Gaps and Opportunities for Academic

Libraries," *Journal of Librarianship and Scholarly Communication* 6, no. 1 (2018), http://doi.org/10.7710/2162-3309.2198.

4. Lisa R. Johnston, Jake Carlson, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert Olendorf, and Claire Stewart, *DCN Team Kick-off Meeting Outputs: Vision, Barriers, Metrics, and Strategic Directions*, July 2016, retrieved from University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/188637.

**To cite this article:** Lisa R. Johnston. "Barriers to Collaboration: Lessons Learned from the Data Curation Network." *Research Library Issues*, no. 296 (2018): 37–43. https://doi.org/10.29242/rli.296.5.