The DataQ Story

by Andrew Johnson

EDITOR'S SUMMARY

Longstanding frustration about the single-purpose use of valuable data led author Andrew Johnson to a career in librarianship, a focus on scientific data curation and, ultimately, creation with colleague Megan Bresnahan of the DataQ project. Johnson's concerns about research data management being a silo within the profession were overcome with the realization that many librarians, especially those at smaller institutions, saw the need for effective curation but felt untrained and unequipped to deal with it. DataQ was established as a service and resource for library personnel to ask questions on data curation and learn through expanding group knowledge. Launched in August 2015, the project draws on a team of 15 editors and additional support members and has been widely and positively received. Funding by the Institute of Museum and Library Services, the Greater Western Library Alliance and others will help expand the effort to provide an important service to the library community.

KEYWORDS

data curation

library technical services

technical support

information science education

continuing education

librarianship

In 2015 Megan Bresnahan and I developed and launched the DataQ project with a vast amount of help from a team of exceptional library professionals from over 30 institutions [1]. Instead of simply giving an overview of the project here, as might be expected, I want to use this opportunity to focus more on the story behind DataQ in the hope that it might inform the larger ongoing conversation about the role of libraries in research data services (and vice versa).

For me personally, the story begins in the middle of the last decade, well before I started my current position as research data librarian at the University of Colorado Boulder in 2011. Back then, I spent most of my time collecting and managing data for a neuropsychology lab at the University of California, San Francisco. The study that I primarily worked on involved data from neuropsychological and neurological tests in the form of paper questionnaires, spreadsheet files and audio and video cassettes. We also collected physical samples (blood and saliva) as well as MRI brain scans. I moved on from this short-lived career after a couple of years, but what never left me was a sense of frustration that all of the rich, complex and valuable data that we had worked so hard to collect would only be used as the basis for a handful of articles before being destroyed following the requisite data retention period. Even more

Andrew Johnson is research data librarian at the University of Colorado Boulder and project lead for DataQ. He can be reached at andrew.m.johnson<at>colorado.edu/.

troubling was the realization that the study participants, all of whom suffered from the same neurological disease, would be asked again and again to donate their valuable time to additional studies. They already provided much of the data to us, which could have been reused had the will and means for widespread data sharing existed.

These were the same frustrations that, on a bit of a hunch, drew me to the field of librarianship in 2009. Since these were the days before even the infamous National Science Foundation data management plan requirements of 2011, I really only suspected that open sharing of research data might be something that academic libraries would be working to promote based on what I knew about the history and values of the profession [2]. On my very first day of graduate school at the University of Wisconsin-Madison School of Library and Information Studies, my suspicion was confirmed when one of my professors mentioned a new informal track being offered in scientific data curation. While I jumped at the chance to pursue solutions to the problems I had seen as a researcher, I worried that I might be pigeonholing myself into what might at the time be a small and unproven niche within the profession. These worries were only reinforced when I started my first professional position in 2011 and quickly realized that not everyone viewed research data management, and specifically data sharing, as a central mission of our profession. Once again, I feared I had found myself in a silo that would always be seen as separate from the rest of the academic library world.

CONTENTS

< PREVIOUS PAGE   NEXT PAGE >

The Vision for DataQ

In an effort to avoid such a siloed fate, I turned to the liaison librarians at my new institution for help. I hoped to find a way to tie what these librarians were already doing in their work to my new role as research data librarian. Working closely with one colleague in particular (Megan Bresnahan, now at Tufts University), we identified a number of liaison librarians who did in fact feel that research data management and sharing were of great importance to their work, but these individuals also expressed some concerns. Namely, they had anxiety about working with research data, and also wanted immediately applicable training and tools to help them engage with researchers about their data needs [3]. We began trying to address these concerns at our institution by developing local training resources, but we also kept an eye toward expanding our work to the wider academic library community [4]. The idea for DataQ came directly from our discussions about the latter.

Megan and I envisioned DataQ as a service and resource that library personnel at any institution could look to when they had questions about research data management and related topics. These could be actual questions they received from researchers or questions of their own that they did not feel comfortable asking on listservs or the like. The questions posed to the DataQ site would then be answered by a team of experts, and all of the questions and answers would be posted for the community to use as a continually evolving resource. Our hope was that DataQ not only would solve users' immediate needs, such as providing answers to their questions, but also would help to reduce the barrier of entry for librarians engaging with research data in their work. As we had attempted to do at our own institution, we wanted DataQ to contribute to the expansion of the community of library personnel working with researchers and their data. We were especially interested in targeting individuals at smaller institutions and/or those who were just starting to work with research data.

Developing the DataQ Service

In order to develop the DataQ service, we applied for and received an Institute of Museum and Library Services Sparks! Ignition Grant for Libraries in 2014, which provided the first year of seed funding for the project. This funding was primarily used to develop the DataQ website and to provide stipends for the expert editorial team whose members would answer the questions posted to the site. Before launching the service, we convened an in-person daylong meeting of the 15 editors, all of whom were selected from a highly competitive pool of applicants, where we brainstormed workflows, ideas for site content and ways that DataQ could reach and support the wider community. The editors identified several significant issues that we had initially overlooked, including functionality that would allow visitors to the site to ask questions anonymously or to identify themselves only to the DataQ project team. This insight proved invaluable since we later found that the vast majority of users of the site chose to remain completely anonymous or to identify themselves only to the editors who were helping to answer their questions. In addition to the pre-launch support that the editors provided, a group of individuals who wanted to help with the project but were unable to join our editorial team, for whatever reason, helped to gather and create initial content for the site launch. A complete list of everyone who has worked on the project to date can be found here: http://ResearchDataQ.org/about.

DataQ Launch and Future Plans

We launched the DataQ site in August 2015 to a positive reception on social media and other online venues. This reception (along with our marketing efforts) helped the site receive close to 18,000 visits from approximately 4,000 users worldwide as of April 2016. During this same period, the editorial team fielded 82 questions in a number of categories, including data management planning, data citation and data sharing, among others. Since our initial grant period ended in late 2015, the Greater Western Library Alliance (an early project supporter) has generously provided funding to continue operating the service, and we are currently in discussions with other national library organizations about plans for long-term sustainability and expansion. The current priorities for DataQ include finding new ways to reach our target audience and improving the site based on current user feedback.

With research data management being discussed widely among funders, publishers, IT professionals and researchers from a number of disciplines, having a strong voice of advocacy for research data access, preservation and reuse (all of which are, of course, enabled by things like good data documentation) remains more important than ever. I thought I heard this voice coming from librarianship when I was in search of a profession with the history and values necessary to tackle the issues that gnawed at me during my brief career as a researcher, but I still worry that this voice is only one of a relative few rather than anything close to that of the mainstream. I hope DataQ, along with a number of other likeminded initiatives like the Research Data Alliance's 23 Things: Libraries for Research Data [5]and the Association of College & Research Libraries' upcoming research data management curriculum [6](to name just two), will continue to build a robust community within libraries that will be able to use the strengths of our profession to fill an important role in the wider research data management conversation.

Acknowledgments

Resources Mentioned in the Article

[1] DataQ: http://researchdataq.org

[2] National Science Foundation. Dissemination and Sharing of Research Results. Retrieved from www.nsf.gov/bfa/dias/policy/dmp.jsp

[3] Bresnahan, M., & Johnson, A. (2013). Assessing scholarly communication and research data training needs. *Reference Services Review, 41*(3).

[4] Johnson, A. & Bresnahan, M. (2015). DataDay!: Designing and assessing a research data workshop for subject librarians. *Journal of Librarianship and Scholarly Communication, 3*(2).

[5] Research Data Alliance. 23 Things: Libraries for Research Data. Retrieved from https://rd-alliance.org/23-things-libraries-research-data-rdas-libraries-research-data-interest-group.html

[6] Malenfant, K. (2015). ACRL selects new curriculum designer/presenters. ACRL Insider. Retrieved from www.acrl.ala.org/acrlinsider/archives/11052