

RESEARCHER APPROVED: A MULTI-INSTITUTIONAL SURVEY OF DEPOSITORS TO SIX ACADEMIC DATA REPOSITORIES

Sarah Wright
sjw256@cornell.edu

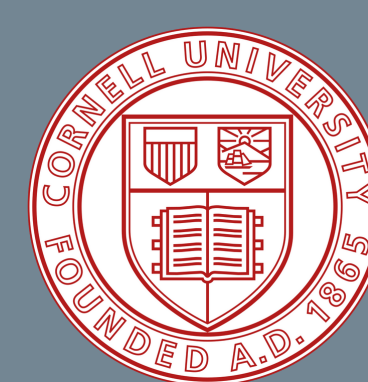
Wanda Marsolek
mars0215@umn.edu

Hoa Luong
hluong2@illinois.edu

Susan Braxton
braxton@illinois.edu

Sophia Lafferty-Hess
sophia.lafferty.hess@duke.edu

Jake Carlson
jakecare@umich.edu



DATA CURATION NETWORK

“ Perspective! When submitting our own datasets, we have a limited view into the repository as a whole, which restricts our own ability to standardize and/or provide information that makes our data easier to find and use by the community. Curators, who see the whole repository as a unified entity, provide that critical perspective. ”
-quote from participant

“ Like a good editor or reviewer on a manuscript, the curation provided by the [name repo] helped to clarify important points about my data that should make it easier to understand when others download it. Small things like units and "what does NA mean here?" likely clear up any confusion, and it means that I get fewer emails asking about small aspects of the data. ”
-quote from participant

Introduction 01

This survey was a collaborative project among six institutions within the Data Curation Network (DCN). The DCN is a network of curators that uses a multi-institutional shared staffing model and standardized CURATED steps [1] to curate data consistently. The DCN explored the value of curation to repository staff (including directors and curators) in a previous survey [2]. The focus of this survey was end-user perceptions of data curation services provided by repositories.



Research question: what are the perceived benefits and impact of data curation work to the researchers who deposited data in our repositories?

References:

- [1] Data Curation Network. CURATED Steps. <https://z.umn.edu/CURATEDSteps>
- [2] Johnston, Lisa et al. (2021). Value of Curation Survey, January 2021. <https://doi.org/10.13020/04ee-q089>.

Methodology 02

Survey method, distributed through Qualtrics



Survey goals:

- to assess user experience and satisfaction with the data curation service provided.
- to understand users' perspective regarding data sharing and the value of the data curation services received.

Survey targets:

- local researchers,
- deposited a dataset with their respective institutional repository between January 1, 2019 and March 15, 2021.

Survey timeline and result:

- opened 2 weeks, between April-June 2021.
- sent 568 emails, received 239 responses (42%).

“ Having another set of eyes review with the idea of sharing and reuse in mind. ”
-quote from participant

“ The suggestions for how documentation should be structured for long term documentation from non-biased outside personnel is the most helpful part. In addition, when I write documentation as the researcher that works with this data every day it is very easy to assume other people know exactly what I am talking about and having someone that has a different background/education look over the submission helps us find where things are unclear. ”
-quote from participant

“ Instructions on how to make metadata more accessible to general audience ”
-quote from participant

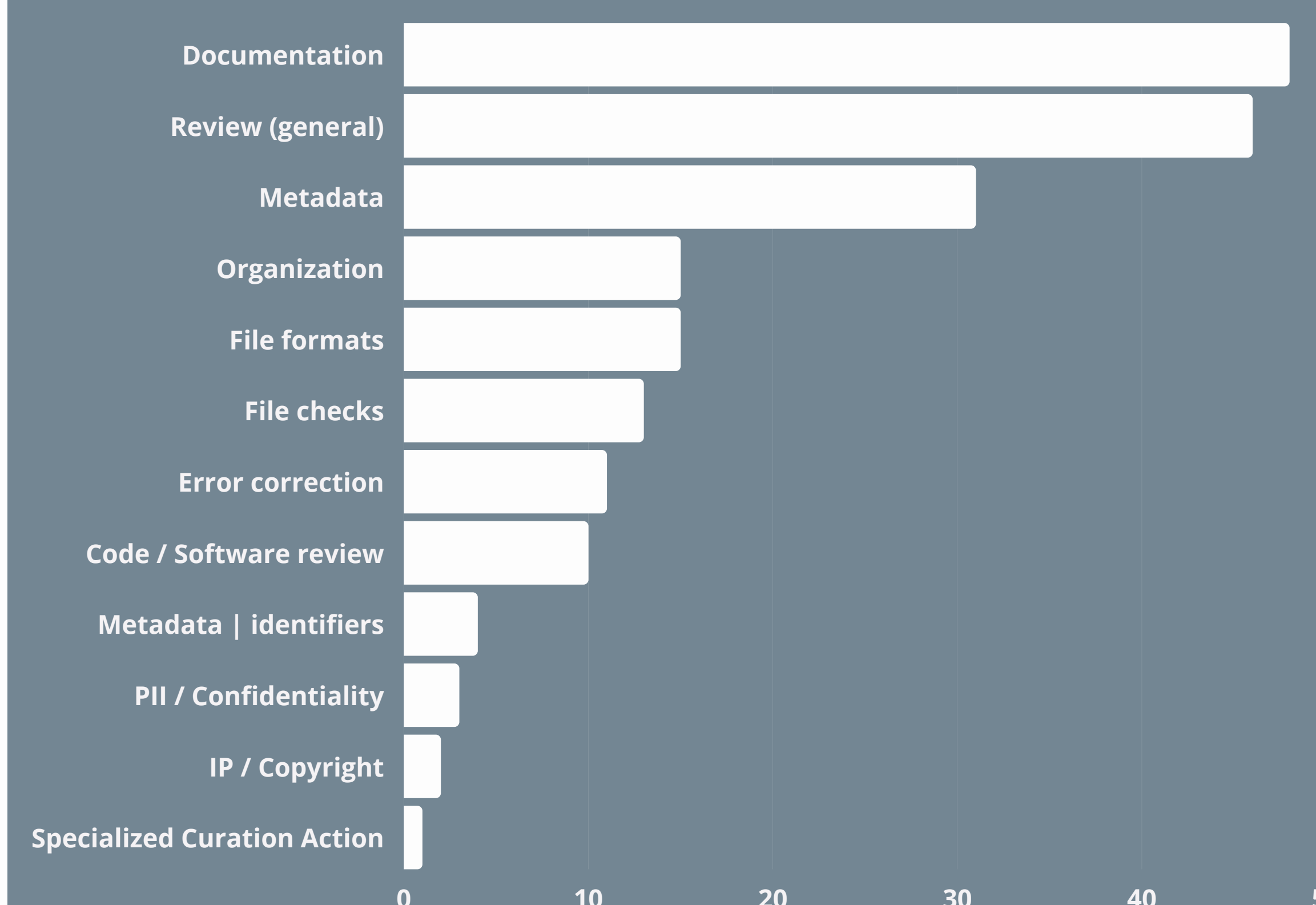
Results 03

- 68%** participants said they expect their local repository staff to curate their dataset
- 98%** agreed that they were satisfied with the review that their dataset received
- 75%** reported changes had been made to their dataset as the result of the curatorial review
- 90%** agreed that they felt more confident sharing their data due to the curation process
- 95%** agreed the impact of data curation has on the data sharing process is worth the effort



Qualitative analysis 04

- We asked participants to identify the most "value-added" curation action taken by the repository.



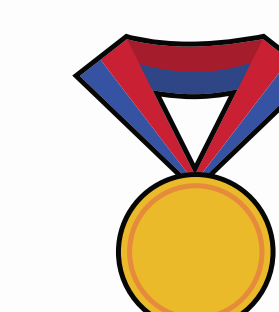
- Curation outcomes were also frequently mentioned, particularly improvements to the data's understandability, access and enabling its use by others.
- Participants also mentioned engagement with curators as being important to them.

Conclusion 05

- Participants were happy to receive curation services.
- Participants reported real improvements to their datasets including greater understandability, access, and potential for re-use though curation actions such as enhanced documentation and enriched metadata.
- Data curation increased participants' confidence when sharing data.
- Data curation is worth the effort.

Acknowledgment:

This research is supported by a grant from the Alfred P Sloan Foundation. Thanks to individuals in the Data Curation Network who contributed feedback and survey testing.



Data Availability statement:

Results and survey instrument are available at the Data Repository for the University of Minnesota, <https://doi.org/10.13020/DZQP-KS53>

