

Opinion

Using collective curation to pay data forward in the life cycle

Ixchel M. Faniel^{1,*}¹OCLC Research, Dublin, OH, USA*Correspondence: fanieli@oclc.org<https://doi.org/10.1016/j.patter.2021.100282>

Drawing from a study of archaeological excavation teams, four collective curation opportunities are proposed to identify and resolve differences in data and documentation practices that arise in team-based research. To create more integrated, well-documented data, the opportunities attend to integrating people rather than technology. The actions people take as data move through the life cycle become the focal point of change.

Several colleagues and I are investigating how data are documented and managed during archaeological research excavations.¹ In a recent article, we proposed collective curation to help archaeological excavation teams more effectively link and contextualize the data being created during their projects.² Although we focused on archaeological excavations, the collective curation opportunities we proposed can be applied to many disciplines where team-based research is being conducted, whether in the field or lab.

Collective curation focuses on integrating people and the actions they take as data move through the life cycle. It provides a way for people to pay data forward. Discussing data and documentation needs at the point of creation provides openings to change that are more likely to be mutually supportive. By identifying and resolving potential issues as they arise, data and documentation loss can be minimized and a more integrated and well-documented set of data can emerge from the project.

“Collective curation provides an opportunity for the team to discuss, negotiate, and form a consensus about what data and documentation are important, given different research objectives and data practices. Collectively curating data at the point of creation also provides a level of quality control. The team is more likely to identify points of failure where data cannot be linked or are not described well enough to be effectively evaluated for use or reuse.”²

As the leaders of archaeological excavations, project directors work with field supervisors who are responsible for excavating archaeological sites and documenting excavation procedures and relationships among the excavated materials. They also work with specialists who contribute to the interpretations drawn about archaeological sites by analyzing the materials and creating data and documentation in the process. Ideally, project directors should be able to link and contextualize specialist data with excavation data so others can understand and trust specialists' analyses and reuse the data. However, this rarely was the case for the four projects we studied.

Wanting to be good collaborators, project directors were hesitant to provide specialists guidance about the data and documentation they needed. Their hands-off approach resulted in data silos. This undoubtedly has negative impacts on data reuse, but it also negatively impacted data use within the team. Our approach to resolve these tensions was to identify opportunities for collective curation that balanced the project directors' desire to maintain collaborative relationships with their need to have specialist data managed and documented so it could be integrated with excavation data.

Opportunities for collective curation start with documentation

Specialists conduct their studies in different ways given their research interests. How they select samples and how they produce and analyze data are critical to ensuring their intellectual independence and integrity are maintained. Given

these differences, specialists' clear and consistent documentation of procedures and rationale are all the more critical. With this documentation, the conclusions they draw from their data analyses are more likely to be understood and trusted. We suggested forming a small group including key data producers on excavation teams to establish a shared set of written guidelines for data documentation as an opportunity for collective curation. Collaboration among those most impacted by the guidelines allows them to learn from each other what is important to document and why. It also creates the buy-in necessary to improve documentation quality and consistency.

Opportunities for collective curation rely on negotiation of needs and expectations

Project directors do not talk about their data and documentation needs with specialists, because specialists' skills and preferences vary and they often are independent researchers not paid by the project. Although project directors are able to get their research done without providing data specifications to specialists, it is not without challenges. Additional time and effort go into decoding specialist data, performing manual data entry, or tracing and adding provenance information to link and contextualize the data. We suggested project director-specialist meetings as an opportunity for collective curation. Discussing what data will be produced and when data will be available puts both in a better position to identify and negotiate needs. We designed a



specialist agreement to guide the conversation with the goal being a joint data curation plan that supports sharing data with each other, making data reusable for others, and negotiating a set of expectations around terms of use.³

Opportunities for collective curation share knowledge to minimize structural silos

Specialists create data to share with the project director and include in their own ongoing research. More often than not, the requirements are at odds. Depending on their expertise and methodological approaches, specialists create and store different types of data in different formats. The data systems, structures, and identifiers project directors use to store, link, and contextualize excavation data also vary given preferences and staff expertise and support. We suggested sharing knowledge about each other's data environment as an opportunity for collective curation. This kind of joint learning can help develop the understanding needed to be open to adjusting data practices in ways that support the needs of the other. Only after understanding each other's processes and practices can the collaborative work to redesign workflows and update data entry systems be done to establish the necessary paths needed to effectively integrate data.

Opportunities for collective curation require intellectual as well as data integration

As subject matter experts, specialists are mainly responsible for the analysis of the materials being excavated. Left to work independently, often in a different location and at a different time than the excavation, real-time use of their findings can

be limited. With autonomy comes fewer occasions for specialists to interact with the team, particularly those excavating the materials who have the potential to contribute to conclusions specialists draw during data analysis. We encouraged formal and informal conversations between specialists and key members of the project team responsible for excavation as an opportunity for collective curation. Providing time and space for both to share what they are learning and experiencing as the excavation unfolds can lead to real-time changes that help to inform the excavation. In addition, these conversations can help shape the research questions project directors and specialists pose, their data collection plans, and the conclusions specialists draw from the analyses they perform.

Conclusion

The opportunities for collective curation we identified aim to balance the tensions project directors experience in their need to be both a collaborator and director. These tensions are not uncommon. Researchers responsible for managing and sharing team-based data in other disciplines face similar issues, as do data repository staff, museum curators, and others participating in the data life cycle, which is why collective curation is so important. The activities performed over the life of data are interconnected. The actions people take to create, document, manage, curate, share, and reuse data can have positive as well as negative impacts on others who also are working with the data. Having more eyes on the data and documentation as they are being created helps identify and resolve failings before they become irreversible. With a primary focus on integrating people rather

than technology, the actions people take as data move through the life cycle become the focal point for change. With each opportunity for collective curation we identified, project directors achieved some level of success in creating a more fully integrated, contextualized set of project data that could be used more effectively within the team, which is an important step toward creating more reusable data for others outside of the team.

ACKNOWLEDGMENTS

The Secret Life of Data project was made possible by a grant from the National Endowment for the Humanities (grant no. PR-234235-16). Any views, findings, conclusions, or recommendations expressed in this article do not necessarily represent those of the National Endowment for the Humanities.

REFERENCES

1. The Alexandria Archive Institute. The Secret Life of Data, <https://alexandriarchive.org/secret-life-of-data/>.
2. Faniel, I.M., Austin, A., Kansa, S.W., Kansa, E., Jacobs, J., and France, P. (2021). Identifying Opportunities for Collective Curation During Archaeological Excavations. *Int. J. Digit. Curation* 16. Published online April 18, 2021. <https://doi.org/10.2218/ijdc.v16i1.742>.
3. Kansa, S.W., Austin, A., Faniel, I.M., and Kansa, E.C. (2020). Specialist Agreement Template. Zenodo. <https://doi.org/10.5281/zenodo.4346486> <https://zenodo.org/record/4346486#.YKOUkpBKg2w>.

About the author

Ixchel M. Faniel, PhD, is a senior research scientist at OCLC. Her interests include improving how people discover, access, and use and reuse content. Her current research examines how academics manage, share, and reuse research data and librarians' experiences designing and delivering supportive research data services. Her research has been funded by the National Science Foundation, Institute of Museum and Library Services, and National Endowment for the Humanities. Her ORCID is <http://orcid.org/0000-0001-7302-5936>.