Conversations in Conservation:
A Data Curation Profile on the Field of Conservation Science

Presented by:
Alice Fracchiolla, Qiqi Lin, Ava Lohr, and Claire Stephens

IS 262A: Data Management and Practice
Prof. Alicia Mara

University of California, Los Angeles
Department of Information Studies
Fall 2025

Introduction

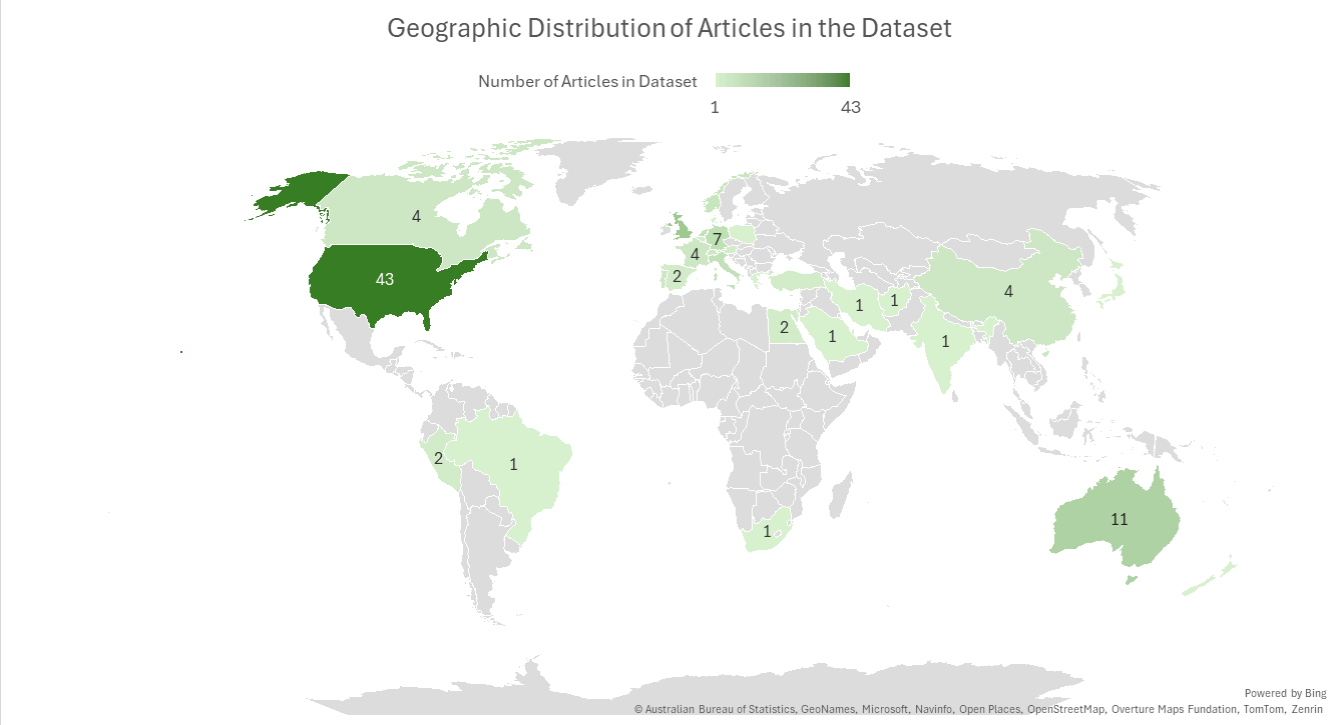
The field of cultural heritage conservation is intimately concerned with the preservation, caretaking, and maintenance of access to human material culture and history. Researchers work directly with objects and the systems and institutions that house them to determine the best processes, protocols, and practices for preserving them through rigorous scientific research. This research often generates data on the types of materials used to create an artwork, their chemical composition, optimal conditions for storing and exhibiting them, and other quantitative data, while contextualizing this data in qualitative and holistic ways. It is easy to see the potential value in opening access to these datasets for the field of conservation more broadly; if one institution makes a breakthrough in, say, how to clean the surface of a painting treated with a particularly fickle varnish, another institution could look at this data and proceed in conserving a painting within their collection with more information and while expending fewer resources, allowing them to devote those resources to other lesser-studied chemical inquiries. However, for a field so invested in preserving this cultural heritage for the benefit of generations to come, the current data culture within the discipline is less robust than we might imagine.

This Data Curation Profile (DCP) synthesizes data practices within the field of cultural heritage conservation based on a coded dataset of research articles published in two journals: *Studies in Conservation* and the *Journal of the American Institute for Conservation*. These articles span object-based conservation science, environmental systems assessments, digital heritage projects, and qualitative research on conservation ethics and user experience. Despite the field's strong global commitment to preserving cultural materials, the dataset reveals a notable absence of standardized data-sharing practices, inconsistent repository usage, and limited adherence to FAIR (Findable, Accessible, Interoperable, Reusable) data principles. Drawing on both the dataset and published literature, including foundational work by Wilkinson et al. (2016), Tsai et al. (2016), Borgman (2015, Yarmey and Baker (2013), and others, this DCP provides an interpretive examination of the field's evolving data infrastructure, challenges, and opportunities.

Types of Research in Cultural Heritage Conservation

Our dataset reflects several dominant paradigms in cultural heritage conservation. The discipline of conservation is global in nature, with researchers around the world seeking to preserve the cultural heritage of their country, institution, or the world more broadly. The geographic locus of the research represented in our dataset is as broad as the scope of the research itself, though the United States accounts for the largest number of articles in the dataset:

Fig. 1 *Locations of researchers and research institutions represented in the dataset.*



The field of conservation is unique in the sense that most of the research conducted within the field is done to garner practical, and often immediately utilizable, information that will directly influence conservation processes and protocols for a specific object, work, or collection. Object-specific research remains the most common, focusing on material analyses of artworks, archaeological objects, artifacts, and architectural fragments. These studies typically examine chemical composition, degradation processes, pigment identification, or evaluate the suitability of conservation materials. Conservators

typically work directly with materials performing chemical, biological, physical, and other forms of scientific research.

A second category includes systems-level or environmental studies, which investigate storage conditions, environmental monitoring, climate change effects, and preventative conservation. They may use institutionally-specific case studies, but are more interested in making recommendations that can be applied across the field. These studies often rely on quantitative modeling, environmental simulations, or geospatial measurements, producing large volumes of data.

Finally, qualitative and mixed-methods research appears throughout the dataset, including studies of conservation ethics, community engagement, user experience, archival workflows, and policy. The frequent presence of this kind of research also distinguishes the field of conservation from other traditional scientific disciplines, blurring the lines between the humanities and so-called “hard” sciences. Intimately linked to the cultural heritage sector institutionally and collegially, the discipline tends to engage with the culturally and historically situated nature of their work more directly than other scientific disciplines. These studies highlight the social and interpretive dimensions of conservation, which Tsai et al. (2016) identify as forms of research often marginalized in data-sharing policy and infrastructure more broadly.

Data Types Generated in the Field

Conservation science generates a diverse range of quantitative, qualitative, and mixed data. Quantitative data can include chemical analyses (e.g., FTIR, XRF, Raman spectroscopy), environmental variables (temperature, humidity, light exposure), degradation rates, radiocarbon dating, and microscopy-based measurements. Three-dimensional modeling, multispectral imaging, and digital reconstructions increasingly contribute to data production. There is a wide array of data types employed by the field, which can contribute to issues of standardization.

Qualitative data, though less likely to be deposited publicly, can include conservation workflow documentation, interview transcripts, user experience surveys, curatorial reflections, community-engaged research notes, and narrative descriptions of treatment processes. These datasets pose challenges for anonymization and long-term standardization, issues identified by Tsai et al. (2016), who argue that data-sharing mandates designed for quantitative disciplines often fail to accommodate qualitative epistemologies.

Journal Landscape and Data Policies

The dataset reflects publication activity across two major conservation journals: *Studies in Conservation* and *Journal of the American Institute for Conservation*. Studies in Conservation is broadly considered the foremost journal in the field, and its scope is international in nature. It typically publishes diverse methodologies in each issue, from more traditional conservation science inquiries to qualitative analyses, and these articles typically emerge from a diverse range of cultural and geographic contexts. The Journal of the American Institute for Conservation (JAIC) is published by the American Institute for the Conservation of Historic and Artistic Works (AIC). As its name suggests, this journal is much more focused on conservation in an American context, and articles most often center on object-specific inquiries.

However, despite their differences, these journals are united in their lack of formal data-sharing requirements. Although both journals encourage authors to follow FAIR principles and reference the FORCE11 Joint Declaration of Data Citation Principles, they stop short of mandating data deposition or recommending trusted digital repositories. Without a clear and mandatory data-sharing policy, and thus incentives for researchers to share their data, data is rarely made available for articles published in these journals. These findings align with Kratz & Strasser (2015), who note that researchers often support the idea of data sharing in theory but fail to deposit data due to time, financial constraints, and lack of

field-specific infrastructure. The field of conservation, which blends scientific and humanistic epistemologies, is particularly vulnerable to these structural gaps.

Repositories Used in Conservation Research

Repository usage in the dataset is inconsistent, with significant variation among authors. Only a small handful of articles deposit data at all, and those that do tend to gravitate towards either their institutional repositories or general-purpose platforms such as Figshare. The figure below summarizes repository usage across the dataset:

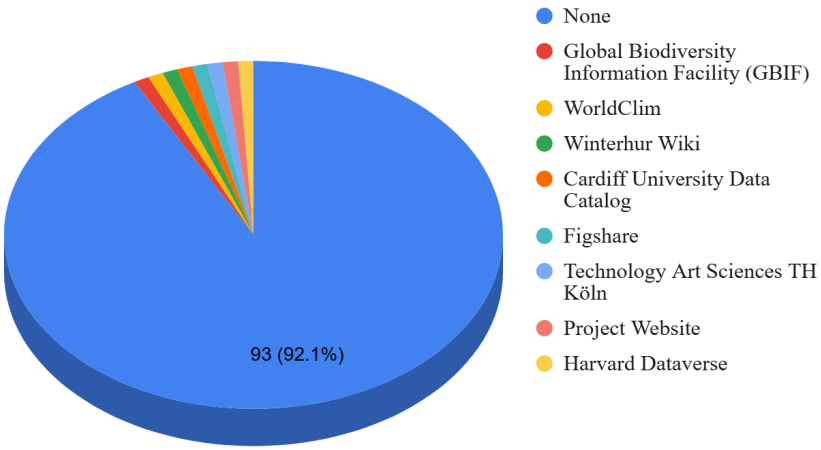


Fig. 2 *Repositories research data were deposited in.*

As the above figure demonstrates, the vast majority of articles did not deposit their data into a designated repository, about 92.1% in fact. In our dataset, there were no recurring repositories used, and each repository listed in the legend only accounts for one data deposit.

Many researchers rely on institutional repositories, personal project websites, or simply no repository at all.These data depositing practices may be unstable or even wholly inadequate for preserving conservation data. Institutional repositories inherently create information silos, and the more researchers deposit their data exclusively in their institutional repositories, the less demand there is for discipline-wide and field-specific repositories that can collate research data from around the world. Rather

than querying a single repository for the data, a researcher would have to find relevant data by seeking it out within each individual repository, limiting discoverability and thus reuse. Personal and/or project-specific websites are created and maintained by researchers themselves, and while they may be able to present data in less strict modes than a traditional repository, the infrastructure of these personalized data deposits may be fragile and need constant maintenance to ensure accessibility, which researchers themselves may not be adequately positioned to keep up with. Additionally, the variance in repository suggests a lack of consensus around data archiving best practices within the field. This concern is echoed by Yarmey and Barker (2013) who emphasize that community standards and interoperable repositories are critical for long-term stewardship.

Notably absent is widespread adoption of the Heritage Science Data Service (HSDS), a UK-based trusted digital repository designed specifically for conservation data. Despite its alignment with FAIR principles and its status as a reputable repository managed by the Archaeology Data Service, none of the over one-hundred articles we collected data from deposited their data here. Granted, HSDS is oriented around a European model of State-funded conservation research, and thus its data library consists primarily of grant-mandated data deposits from researchers within the UK. However, its infrastructural subject-specific design could make it a compelling repository for discipline-wide data storage and circulation.

Funders and Policy Ecosystems

Funding structures in conservation play a critical role in shaping data practices. Unlike biomedical or environmental sciences, where public funders increasingly require data management plans and open data deposition (Wilkinson et al. 2016), cultural heritage funding often comes from museums, private foundations, or specialized cultural agencies within the government whose policies rarely mandate data sharing. This divergence highlights a key structural issue: conservation research is often highly publicly valuable, but it is rare for data generated in the field to be publicly available. Funders such as the

National Endowment for the Humanities, the Getty Foundation, or governmental cultural ministries typically support conservation projects without imposing open-data requirements analogous to those in the natural sciences. The result is a policy environment in which data preservation is *encouraged*, but not enforced.

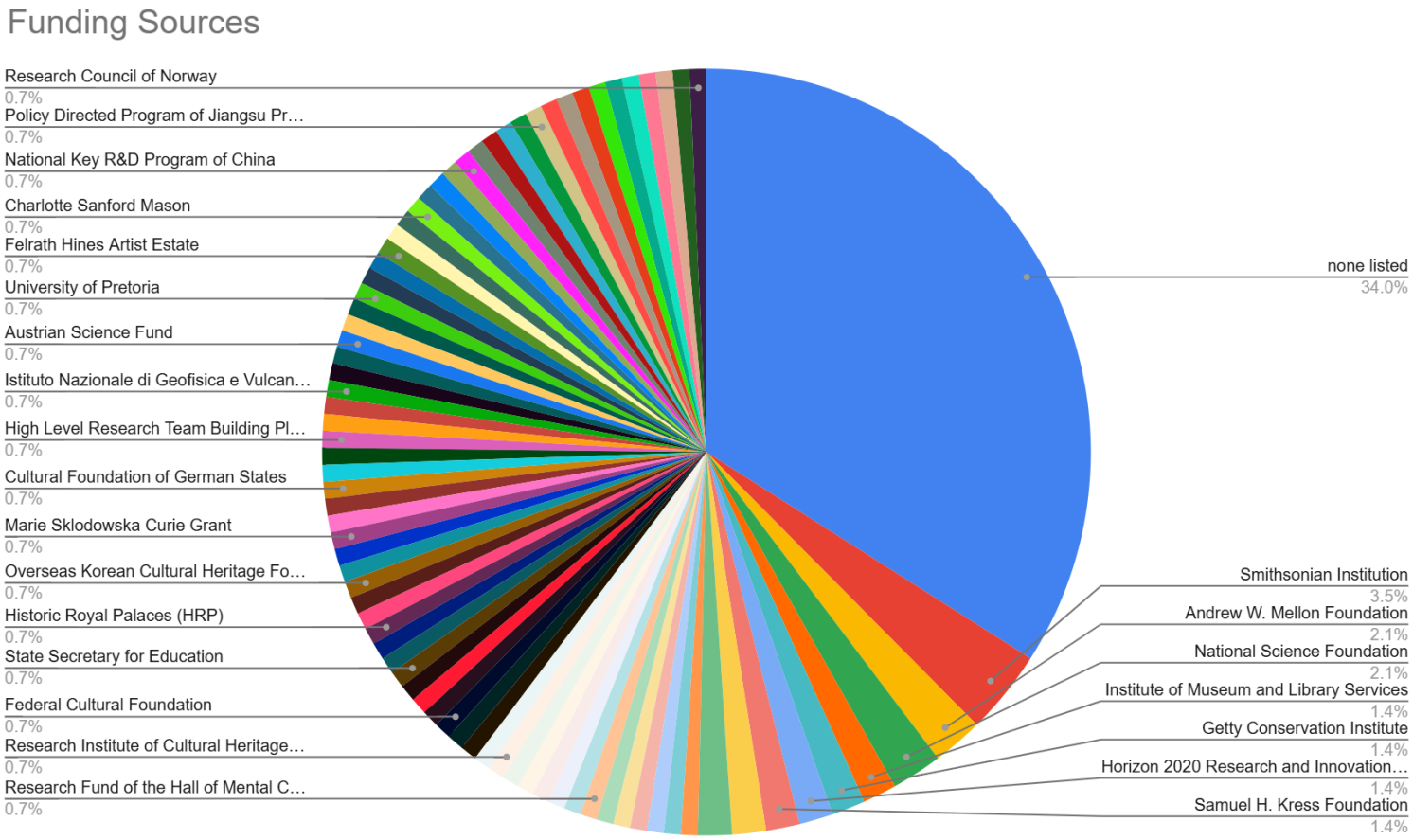


Fig 3. *Distribution of Funding Sources. Most research is funded by highly specialized local agencies, however, there are several recurring funding sources that appear in the dataset, which are often private foundations and cultural institutions in the United States.*

Metadata Standards in the Field

Metadata practices in the dataset are inconsistent and often minimal, reflecting the absence of unified metadata standards across conservation research. While some digital heritage projects employ

Dublin Core, VRA Core or the Lightweight Information Describing Objects (LIDO) schema, scientific datasets may rely on laboratory-specific metadata or no standardized schema at all. This fragmentation poses challenges for interoperability and long-term reuse. FAIR-aligned metadata schemas exist, but their adoption appears limited (Soares et al. 2024). Discipline-specific infrastructure like HSDS may be best positioned to be leaders in the field to encourage the widespread adoption of these FAIR-aligned schemas. However, without community-wide agreement, metadata remains a significant barrier to machine-readable, usable conservation data.

Data Sharing, Availability, and Reproducibility Practices

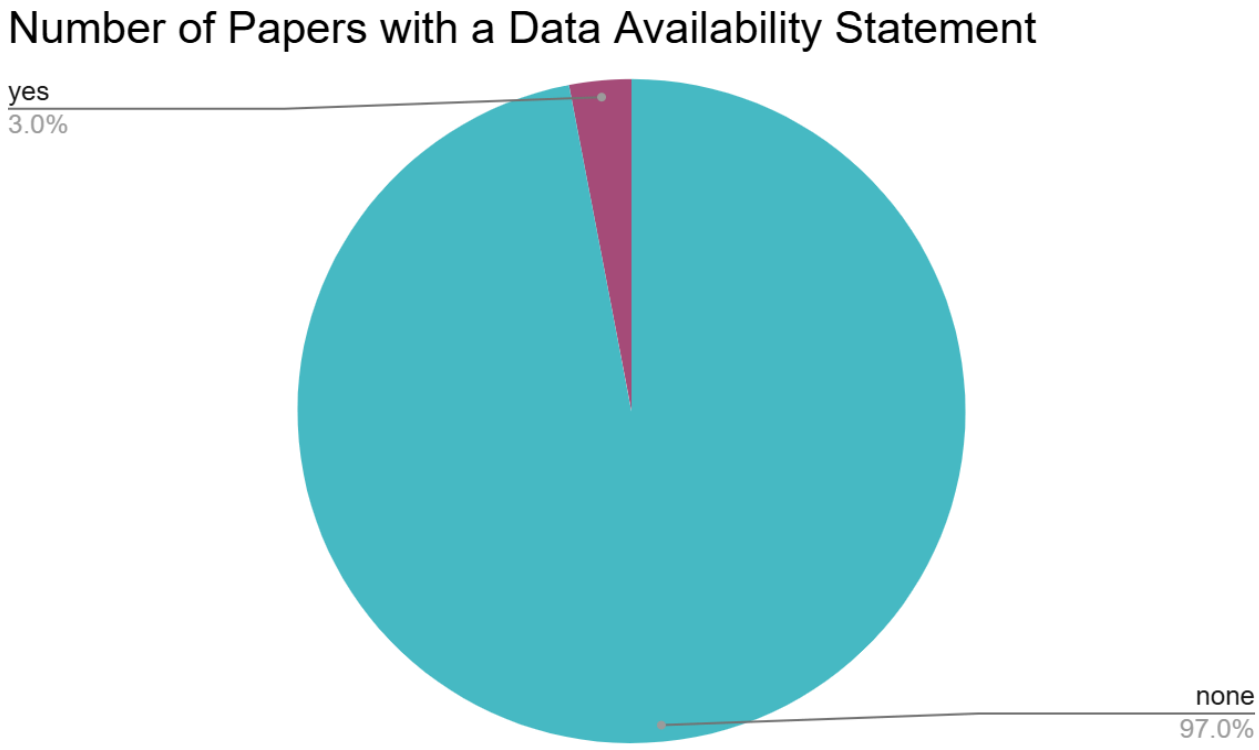


Fig. 4 *Chart showing number of papers that include a data availability statement in the dataset*

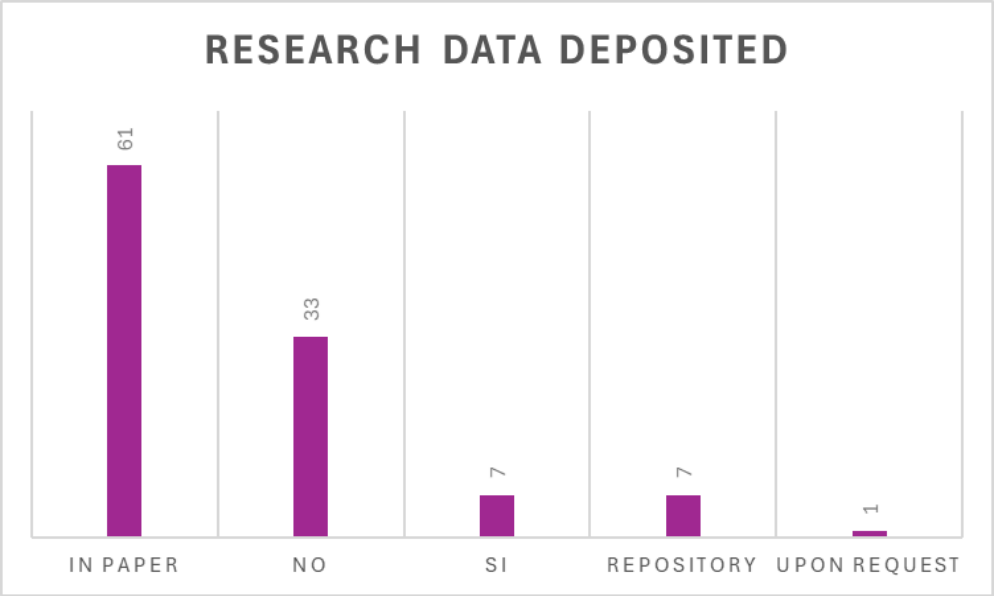


Fig 5. *Graph showing where data was deposited, if at all.*

The dataset clearly shows how few articles include Data Availability Statements, and even fewer actually provide access to raw or unprocessed data. Most of the data that is deposited at all is deposited within papers in the form of charts, tables, or other figures that have transformed the data in some manner, which may have limited utility for future researchers. This limited disclosure may stem from the object-specific nature of many conservation studies. Artifacts being investigated by conservators are typically handled by a single institution, and this may lead researchers to perceive their data as non-reproducible or non-generalizable. However, such assumptions can inhibit comparative research, cross-institutional collaboration, and long-term digital preservation.

Consistent with Borgman's (2015) assertion that data exist within complex knowledge infrastructures, reproducibility in conservation is shaped by institutional context, material constraints, and technological access. Without trusted repositories, clear policies, and consistent metadata standards, reproducibility remains limited even when researchers are theoretically supportive of open science.

Conclusion

For a field so intimately concerned with the preservation of cultural heritage to ensure its discoverability, accessibility, and continued relevance in contemporary life and academic inquiry, the lack of a robust data sharing culture within conservation science seems contradictory to its foundational goals. The dataset, when paired with current literature on best practices in data stewardship, reveals a field in transition. Cultural heritage conservation research continues to expand in scope, methodological diversity, and global reach, yet its data-sharing infrastructure lags behind. Challenges include inadequate repository adoption, inconsistent metadata standards, limited policy enforcement, and structural biases favoring quantitative scientific data over qualitative or mixed methods. While some have drawn a sharp distinction between conservation science and "traditional" scientific disciplines, this distinction is slowly breaking down as the field takes up broader questions, such as climate change, that will affect not only preservation of art and archaeological objects but the preservation of human life more broadly.

As conservation increasingly embraces digital tools, environmental modeling, and large-scale collaborative projects, the need for robust, FAIR-aligned data stewardship becomes more urgent. Perhaps there is a need for new globally-oriented conservation data repositories to emerge and lead the charge in creating a robust conservation data sharing culture, or perhaps it's simply necessary to increase the visibility and advocate for the importance of using existing repositories such as HSDS. Regardless, the onus of responsibility for this data culture is not solely on researchers or their granting institutions; data storage infrastructure will need to catch up to accommodate the unique data conservation science produces in order to maintain and circulate it for generations to come. Strengthening repository infrastructure, adopting interoperable metadata standards, and reforming journal and funder policies will be essential steps toward building a sustainable, ethically grounded data culture in cultural heritage conservation.

Bibliography

Borgman, C. L. (n.d.). *Big Data, Little Data, No Data: Scholarship in the Networked World*. Retrieved December 8, 2025, from https://direct.mit.edu/books/monograph/3085/Big-Data-Little-Data-No-DataScholarship-in-the

Kratz, J. E., & Strasser, C. (2015). Researcher Perspectives on Publication and Peer Review of Data. *PLOS ONE*, *10*(2), e0117619. https://doi.org/10.1371/journal.pone.0117619

Soares, F. M., Pires, L. F., Santos, L. O. B. da S., Calhau, R. F., Maculan, B. C. M. dos S., Coyle, K., Wang, S., Folmer, E., Drucker, D. P., Campos, M. L. de A., Marcondes, C. H., Almeida, M. B., Braghetto, K. R., Dias, G. A., Salim, J. A., Corrêa, F. E., Moreira, D. de A., Delbem, A. C. B., & Saraiva, A. M. (2024). Towards a Conceptual Model for FAIR Metadata Schemas. *Companion Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special Topics, Posters and Demos : Co-Located with ER 2024*, 42–55. https://research.utwente.nl/en/publications/towards-a-conceptual-model-for-fair-metadata-schemas/

Tsai, A. C., Kohrt, B. A., Matthews, L. T., Betancourt, T. S., Lee, J. K., Papachristos, A. V., Weiser, S. D., & Dworkin, S. L. (2016). Promises and pitfalls of data sharing in qualitative research. *Social Science & Medicine*, *169*, 191–198. https://doi.org/10.1016/j.socscimed.2016.08.004

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Yarmey, L., & Baker, K. S. (2013). Towards Standardization: A Participatory Framework for Scientific Standard-Making. *International Journal of Digital Curation*, *8*(1), 157–172. https://doi.org/10.2218/ijdc.v8i1.252