

Fossils in a Digital Age: Preserving 3D Fossil Data in Digital Libraries

Anneliese Eber¹

¹*School of Library and Information Studies, University of Alberta, aeber@ualberta.ca*

To Cite:

Eber, A. (2023). Fossils in a digital age: Preserving 3D fossil data in digital libraries. *Pathfinder: A Canadian Journal for Information Science Students and Early Career Professionals*. 4(1), 223-229. <http://doi.org/10.29173/pathfinder68>

Abstract

Digital libraries are a powerful tool for preserving and disseminating information in a technology-driven world. Providing options for housing non-traditional information, such as three-dimensional (3D) data, digital libraries also address barriers (e.g., geographic location and physical space) that affect traditional libraries. While there are certainly still barriers that may prevent users from accessing and utilizing digital content, digital libraries offer ways of preserving, housing, and disseminating non-traditional information that physical libraries cannot. Within the anthropological community, digital libraries can serve as a method of housing 3D fossil data. There are unique challenges and barriers associated with studying fossils that digital libraries help to break down. Through the preservation and dissemination of 3D fossil data, researchers across the globe can work with fossil data without the concerns associated with handling fragile fossils. However, such 3D digital libraries are not without challenges. While there are clear standards and metadata schemas (e.g., RDA and DCMI) for digital libraries containing traditional materials, no such standardization exists for digital libraries containing 3D data. As such, metadata requirements and elements for existing 3D digital libraries are inconsistent and may lack important details about the data, which can compromise functionality. This paper explores trends and challenges that digital libraries containing 3D data experience and highlights best-practice solutions for future 3D digital libraries to overcome these challenges.

Keywords: digital libraries, three-dimensional data, metadata schemas, data preservation, copyright

Digital libraries are powerful tools for preserving and disseminating information in a technology-driven world. They provide options for storing and preserving non-traditional information, like three-dimensional (3D) data, that is otherwise difficult to access (Skinner et al., 2013). Due to their versatility, digital libraries provide an interdisciplinary solution for storing and disseminating data while breaking down physical barriers to accessing data (e.g., geographic location and physical accessibility

of buildings). One such application of digital libraries is storing 3D fossil scans. While paleoanthropological research has historically used physical fossils, working with fossils can be difficult for a myriad of reasons, including limited access (Hublin, 2013; Skinner et al., 2013). The process of accessing and examining physical fossils can be restricted by the geographical location of fossils, how fossils are accessioned and curated, institutional bureaucracy, and the physical state of fossils (Hublin, 2013; Skinner et al., 2013). To address some of these concerns, virtual paleoanthropology has become increasingly popular with many fossils being sent to scanning facilities to create digital copies; however, the dissemination of these fossil scans remains difficult (Hublin, 2013). For example, MorphoSource from Duke University is a digital repository that holds many anthropological models, but the metadata and browsing features make discoverability difficult for non-experts (Hall et al., 2019). Opposite this, platforms like Sketchfab and Thingiverse support better discoverability for non-experts but lack the metadata to support the research community (Hall et al., 2019). Despite their challenges, the use of digital libraries provides an accessible way to create a repository of 3D data and fossil scans that can be used by different audiences globally.

Trends and Challenges in 3D Digital Libraries

Although digital libraries hold the power to create an easily accessible repository of fossil data, removing geographic barriers, and preserving the integrity of fossils, such a library poses unique problems due to the nature of the data. These problems include how to catalogue, acquire, and preserve the data in addition to how to protect the rights of those who create and own the data (Berquist et al., 2012; Boyer et al., 2017; Davies et al., 2017; Hardesty et al., 2020; Skinner et al., 2013; Ulguim, 2018; Zamyadi et al., 2013). Literature about previous and existing digital fossil libraries, and similar digital libraries containing 3D data, provides important insight into how to create, maintain, and grow a repository of 3D fossil data moving forward.

Metadata Schemas

One of the biggest inconsistencies in current online digital libraries housing 3D data is a lack of standardized metadata schemas both within and between digital libraries (Hardesty et al., 2020). In the past, schemas have used modifications of Dublin Core (Mi & Pollock, 2018) but previous modifications of Dublin Core schemas do not fit

all 3D digital libraries. Due to the vast types of information captured in these varying types of research, previous studies can provide guidance on creating custom metadata schemas; however, typically each digital library must create a unique metadata schema appropriate for the collection and the platform it is housed on (Berquist et al., 2012; Mi & Pollock, 2018). For example, a schema might need to include fields for crowd-sourced keywords if the collection is housed on a platform that supports these functions (Mi and Pollock, 2018).

Regardless of the specific elements that are included in a schema, proper records that support maximum usability should include descriptive, administrative, and structural metadata elements (Joudrey & Taylor, 2017). Additionally, metadata must function to aid discoverability, usability, and interoperability of the stored information (Hall et al., 2019; Mi & Pollock, 2018; Ulguim, 2018). Fields that support searchability and discoverability might include institutional identifiers, holding location information, and keywords (Hardesty et al., 2020; Mi & Pollock, 2018). Similarly, fields containing information about physical characteristics of an object and the scan specifications contribute towards the usability of information (Davies et al., 2017; Hall et al., 2019; Mi & Pollock, 2018). However, partly due to the nature of 3D data, interoperability remains a significant challenge in 3D digital libraries (Albrezzi et al., 2022; Hall et al., 2019). Despite challenges in interoperability, combining these foundational qualities has led to the successful creation of metadata schemas and FAIR (findable, accessible, interoperable, and reusable) data that fully represent and support the function of different types of collections (Golubiewski-Davis et al., 2022; Hall et al., 2019; Mi & Pollock, 2018, Zamyadi et al., 2013). While existing 3D digital libraries provide an excellent starting point for digital fossil libraries, they demonstrate a need for additional discussion and clarity about metadata schemas and metadata implementation within collections.

Data Acquisition and Preservation

Creating a digital fossil library of 3D scans of fossil specimens requires that data be acquired or created in a way that both supports life in a digital library and multi-purpose use of the data (Berquist et al., 2012; Davies et al., 2017; Skinner et al., 2013;

Ulgum, 2018; Zamyadi et al., 2013). While 3D scans of fossils are often produced with a specific type of research in mind, for the purpose of a 3D digital library, scans must be of high enough quality, and in industry-standard file types, to support research that will use data in diverse ways that may diverge from the original intended purpose (Albrezzi et al., 2022; Davies et al., 2017; Hardesty et al., 2020; Skinner et al., 2013). To further support the long-term preservation of the data, all information about data acquisition and creation should be included in the metadata schema and should be mindful about best-practices within the field (Davies et al., 2017; Hall et al., 2019; Hardesty et al., 2020). Additionally, accurate, stable, and long-term identification of data is integral to the longevity of 3D digital libraries (Davies et al., 2017). As such, stable identification numbers such as digital object identifiers (DOIs) or archival resource keys (ARKs) provide a solution to long-term identification and method for citing data (Blundell et al., 2022; Davies et al., 2017).

Copyright Considerations

To further support the long-term access and use of data, copyright of data needs to be addressed. One concern data authors and owners might have when publishing their data to digital repositories is the uncertainty of how it will be used, who will use it, and whether credit will be given to data authors and owners. To address concerns about intellectual property and data ownership, different systems have been proposed to track data usage, who is using the data, and for what research (Berquist et al., 2012; Boyer et al., 2017; Davies et al., 2017; Hublin, 2013; Skinner et al., 2013). These methods range from 'checking out data' as in a traditional library to submitting letters of intent to institutions before access will be granted (Berquist et al. 2012; Skinner et al. 2013). Regardless of how users access data, these methods help to ensure research projects are not being repeated simultaneously which protects the data owners, data authors, and other researchers with a personal stake in the data (Berquist et al. 2012; Skinner et al. 2013).

To address copyright issues, data owners and authors might assign Creative Commons licenses to the data (Boyer et al., 2017; Davies et al., 2017). Boyer et al. (2017) suggest that CC-BY-NC is the preferable license for academic data as it allows secondary users of the data, with attribution, to reuse and modify the data. Davies et al.

(2017) also offer CC-BY-ND as a potential license that might be useful, though, this is a more restrictive license as it does not allow for derivative data to be published therefore limiting academic publications. However, copyright licenses must be assigned by the data owner and therefore must consider who the data authors and data owners are if they are not the same (Boyer et al., 2017). While the copyright system is not perfect, it does allow some security for data authors and data owners who choose to share their data. However, while current digital fossil libraries fail to ensure copyright information is present if it is available at all, copyright information is integral and should be included in all records (Blundell et al., 2022; Boyer et al., 2017; Davies et al., 2017).

Conclusion

Digital collections of 3D data are certainly not a new conception and they have been slowly growing and being built over the last decade (Boyer et al., 2017; Skinner et al., 2013). Despite the potential of 3D digital libraries for the anthropological community, the application of such libraries needs further thought and clarity. In addition to issues concerning copyright and how researchers use and access the data, current metadata schemas are lacking in the foundational metadata elements. While these discussions are lacking as they pertain to digital fossil libraries, previous work on other 3D digital libraries like MorphoSource provides guidance of how such digital fossil libraries can be improved. To increase the discoverability, usability, and interoperability of these collections, current digital libraries should revisit the metadata schemas and core principles to expand upon current schemas and ensure all stored models meet the metadata requirements.

Works Cited

- Albrezzi, F., Bonnett, J., Gniady, T., Richards-Rissetto, H., & Snyder, L. M. (2022). Accessing 3D Data. In J. Moore, A. Rountrey, & H. Scates Kettler (Eds.), *3D data creation to curation: Community standards for 3D data preservation* (pp. 259–295). Association of College and Research Libraries.
- Berquist, R. M., Gledhill, K. M., Peterson, M. W., Doan, A. H., Baxter, G. T., Yopak, K. E., Kang, N., Walker, H. J., Hastings, P. A., & Frank, L. R. (2012). The digital fish

- library: Using MRI to digitize, database, and document the morphological diversity of fish. *PLOS ONE*, 7(4), e34499. <https://doi.org/10.1371/journal.pone.0034499>
- Blundell, J., Clark, J. L., DeVet, K. E., & Hardesty, J. (2022). Metadata requirements for 3D Data. In J. Moore, A. Rountrey, & H. Scates Kettler (Eds.), *3D data creation to curation: Community standards for 3D data preservation* (pp. 157–204). Association of College and Research Libraries.
- Boyer, D. M., Gunnell, G. F., Kaufman, S., & McGeary, T. M. (2016). MorphoSource: Archiving and sharing 3-D digital specimen data. *The Paleontological Society Papers*, 22, 157–181. <https://doi.org/10.1017/scs.2017.13>
- Davies, T. G., Rahman, I. A., Lautenschlager, S., Cunningham, J. A., Asher, R. J., Barrett, P. M., Bates, K. T., Bengtson, S., Benson, R. B. J., Boyer, D. M., Braga, J., Bright, J. A., Claessens, L. P. A. M., Cox, P. G., Dong, X.-P., Evans, A. R., Falkingham, P. L., Friedman, M., Garwood, R. J., ... Donoghue, P. C. J. (2017). Open data and digital morphology. *Proceedings of the Royal Society B: Biological Sciences*, 284(1852), 20170194. <https://doi.org/10.1098/rspb.2017.0194>
- Golubiewski-Davis, K., Maisano, J., McIntosh, M., Moore, J., Niven, K., Rourk, W., & Snyder, R. (2022). Best practices for 3D data preservation. In J. Moore, A. Rountrey, & H. Scates Kettler (Eds.), *3D data creation to curation: Community standards for 3D data preservation* (pp. 15–88). Association of College and Research Libraries.
- Hall, N. F., Hardesty, J., Lischer-Katz, Z., Johnson, J., Cook, M., Griffin, J., Ogier, A., Carlisle, T., Xie, Z., McDonald, R., & Wittenberg, J. (2019). Challenges and directions in 3D and VR data curation. *International Journal of Digital Curation*, 14(1), 150–167. <https://doi.org/10.1371/journal.pone.0034499>
- Hardesty, J. L., Johnson, J., Wittenberg, J., Hall, N., Cook, M., Lischer-Katz, Z., Xie, Z., & McDonald, R. (2020). 3D data repository features, best practices, and implications for preservation models: Findings from a national forum. *College and Research Libraries*, 85(5), 789–807. <https://doi.org/10.5860/crl.81.5.789>
- Hublin, J.-J. (2013). Free digital scans of human fossils. *Nature*, 497(7448). <https://doi.org/10.1038/497183a>

- Joudrey, D. N., & Taylor, A. G. (2017). *The organization of information* (4th edition). Libraries Unlimited.
- Mi, X., & Pollock, B. M. (2018). Metadata schema to facilitate linked data for 3D digital models of cultural heritage collections: A University of South Florida Libraries case study. *Cataloging & Classification Quarterly*, 56(2–3), 273–286. <https://doi.org/10.1080/01639374.2017.1388894>
- Skinner, M. M., Kivell, T. L., Potze, S., & Hublin, J.-J. (2013). Microtomographic archive of fossil hominin specimens from Kromdraai B, South Africa. *Journal of Human Evolution*, 64(5), 434–447. <https://doi.org/10.1016/j.jhevol.2013.01.007>
- Ulgum, P. (2018). Models and metadata: The ethics of sharing bioarchaeological 3D models online. *Archaeologies*, 14(2), 189–228. <https://doi.org/10.1007/s11759-018-9346-x>
- Zamyadi, A., Pouliot, J., & Bédard, Y. (2013). Towards precise metadata-set for discovering 3D geospatial models in geo-portals. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-2-W1, 299–308. <https://doi.org/10.5194/isprsannals-II-2-W1-299-2013>