

The Case for Archiving Legacy Archaeological Collections into tDAR: Digitizing Approaches for Preservation, Dissemination, and Access

DAVID M. PLAZA
Principal Investigator
david.plaza@enmu.edu

Abstract

Paper documents are a time proven technology for the preservation and dissemination of information. The advent of computers and the Internet have enhanced some basic concepts for which paper documents have traditionally been used, such as the management of information on an external platform. Some benefits of using computers are the speed and ease of aggregating and disseminating information. This research will review preservation methods on how to ingest a legacy archaeological collection into an international online digital archive repository called the Digital Archaeological Record (tDAR). This online repository is housed by the Center for Digital Antiquity and is a secured databank that has relational capabilities to combine various datasets from different archaeological investigations. This online archival repository was established to improve and assist in the preservation, dissemination, and access of archaeological data, which is a nonrenewable resource. The archaeological legacy collection used here is the Anasazi Origins Project (AOP). The AOP is particularly significant for its contribution to defining the Archaic northern Southwest. The benefit of ingesting this collection into tDAR is to establish the foundation to centralize various components such as paper documents, artifacts, and photographs, of the AOP collection located across the United States through virtual unification. This research contributes to the ongoing discussion of whether there is a need to establish a professional niche that effectively integrates archaeology, museum studies, archiving, and information technologies. Moreover, this research will explore the possibility of "big data" in archaeology.

The AOP Legacy Archaeological Collection

Cynthia Irwin-Williams designed the Anasazi Origins Project (AOP) and served as the principal Investigator. The objective of this monumental multidisciplinary project was to investigate the antecedents of the Ancestral Puebloans, which she called the Oshara Tradition. This project was initiated at a time when the Archaic period of the northern Southwest of the United States was poorly understood. The result of this project was a cultural historical and processual model to aid in understanding this time period as well as an enormous collection^{1,2}.

The AOP Legacy Archaeological Collection housed at Eastern New Mexico University (ENMU) has significant data potential and is a representative sample of the Archaic northern Southwest. In this research, this collection is treated as a legacy dataset, which is defined as any set of data that is old, unused, obsolete, and/or disparate that have been passed from one generation to the next. The current efforts to maintaining this collection at ENMU are lead by Dr. John Montgomery and, under his supervision, graduate and undergraduate students. Currently, the AOP collection has been organized and is suitable for research. In addition, there is ongoing work to digitize this collection by entering and organizing the artifacts into Excel spreadsheets. Another effort is the Anasazi Origins Project Digital Archives Initiative objective to digitize the papers documents of AOP collection for preservation and access^{1,2}.

tDAR

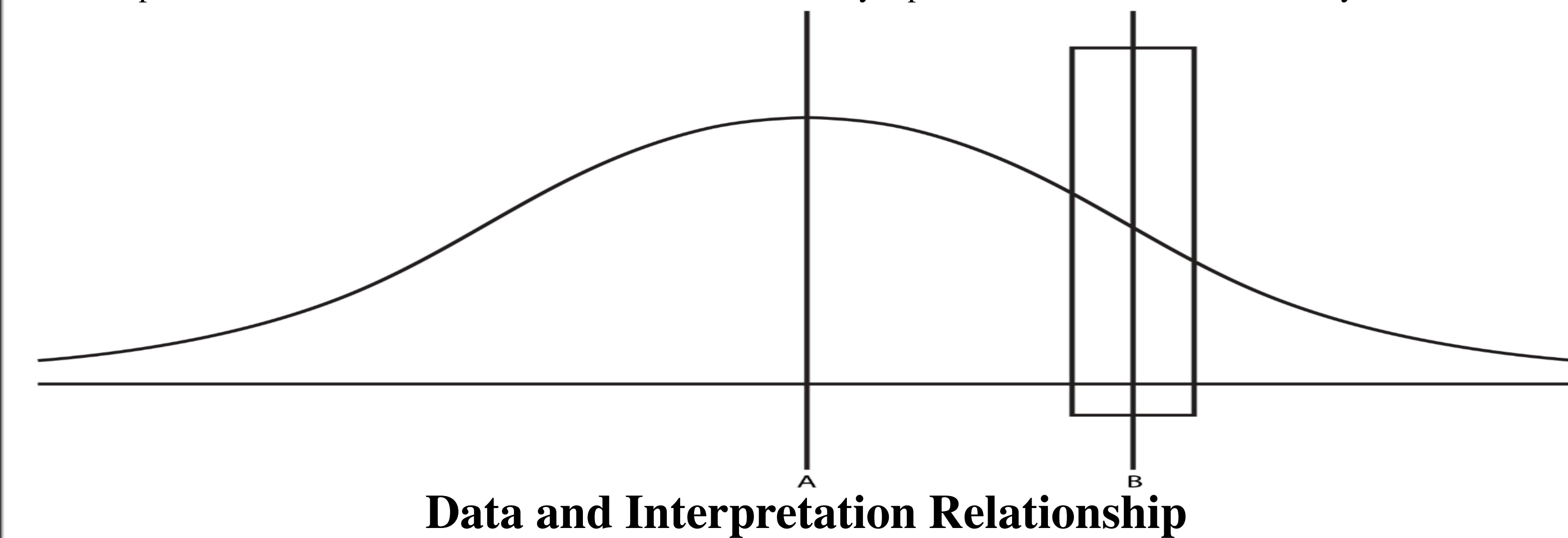
The Digital Archaeological Record (tDAR) is an international online digital repository and archive housed by Digital Antiquity, a nonprofit organization devoted to enhancing preservation and access of irreplaceable archaeological records. The benefits of tDAR are the accessibility and preservation of archaeological datasets from other institutions that are continuously enhanced. The added bonus is that it is a viable starting point for gathering data for research without ever leaving the home institution. Moreover, tDAR offers tools to combined datasets regardless of their ontology³.

Methodology

- **Photocopying** – Site records and inventory sheet for notes photocopied onto acid free paper.
- **Scanning** – Site records, inventory sheet for notes, and publications were saved as a PDF and maps were saved as a JPEG using a scanner. 300 dpi was the minimum items were saved as².
- **Photography** – Some USGS 7.5 minutes series maps were digitized using a SLR camera because of the degraded condition of them due to archaeological field work. This method is a viable option when a record cannot be moved or there is an issue with light affecting the record.
 - Scanning and photograph may not capture the entire document. Digitizing sections of the documents and photomerging them using Adobe Photoshop was a necessary step for some documents².
- **Database Entry** – MS Access was used to create a dataset of the AOP paper documents a to be uploaded onto tDAR. This dataset was organized using the Laboratory of Anthropology site record to help better integrate with existing data from other projects².
- **Uploading onto tDAR** – Create a "project" on tDAR. A project contains and organizes items as resources into child resources that can inherit the metadata of the project. Resources can also be organized into collections to keep relevant items together².
- **Metadata** – is used to create contextual information and make information searchable. Creating metadata consisted of two primary efforts: 1. enter a complete citation so that the resource can be identified and authors credited and 2. tDAR online form to populate a keyword list (location, time period, cultural affiliation, and components of collection) so that users can find items using the search function².

Results and Discussion

Some of the components of the AOP legacy archaeological collection has been ingested into tDAR as a fluid electronic dataset that can be easily and quickly replicated as well as preserved on an alternative repository that is formatted to facilitate analysis. In addition, this project allows this collection to be more easily reintroduced in the current archaeological dialogue. The importance of this project goes beyond preservation or answering specific archaeological questions, it contributes to big data. This type of data is defined here as a compilation of large (ex. terabytes to petabytes) and complex (diverse) datasets that can be used to answer questions. The availability of big data within a respective discipline enables scholars to identify patterns more clearly and to have more accurate interpretations⁴. The continued effort to preserve and centralize this collection on an online international repository that is regulated by an institution dedicated for scholarly as well as ethical work can facilitate the fluidity of this model to other collections to help integrate similar collections into the realm of big data. Datasets should be preserved to enhance our interpretations. Conversely, the use of only a portion of data that is readily available can skew interpretation as shown in the graph below. The bell curve represents all the body of datasets collected from archaeological projects. Datasets closest to the peak of the curve are relevant to answering a particular question. Line A represents the likelihood of accurate interpretation, when all datasets relevant to the research question from previous as well as contemporary projects are available and used. If datasets from previous or contemporary archaeological projects are not used, interpretations can be skewed, as seen with line B, which only represents a subset of the entire body of research¹.



Conclusion and Future Research

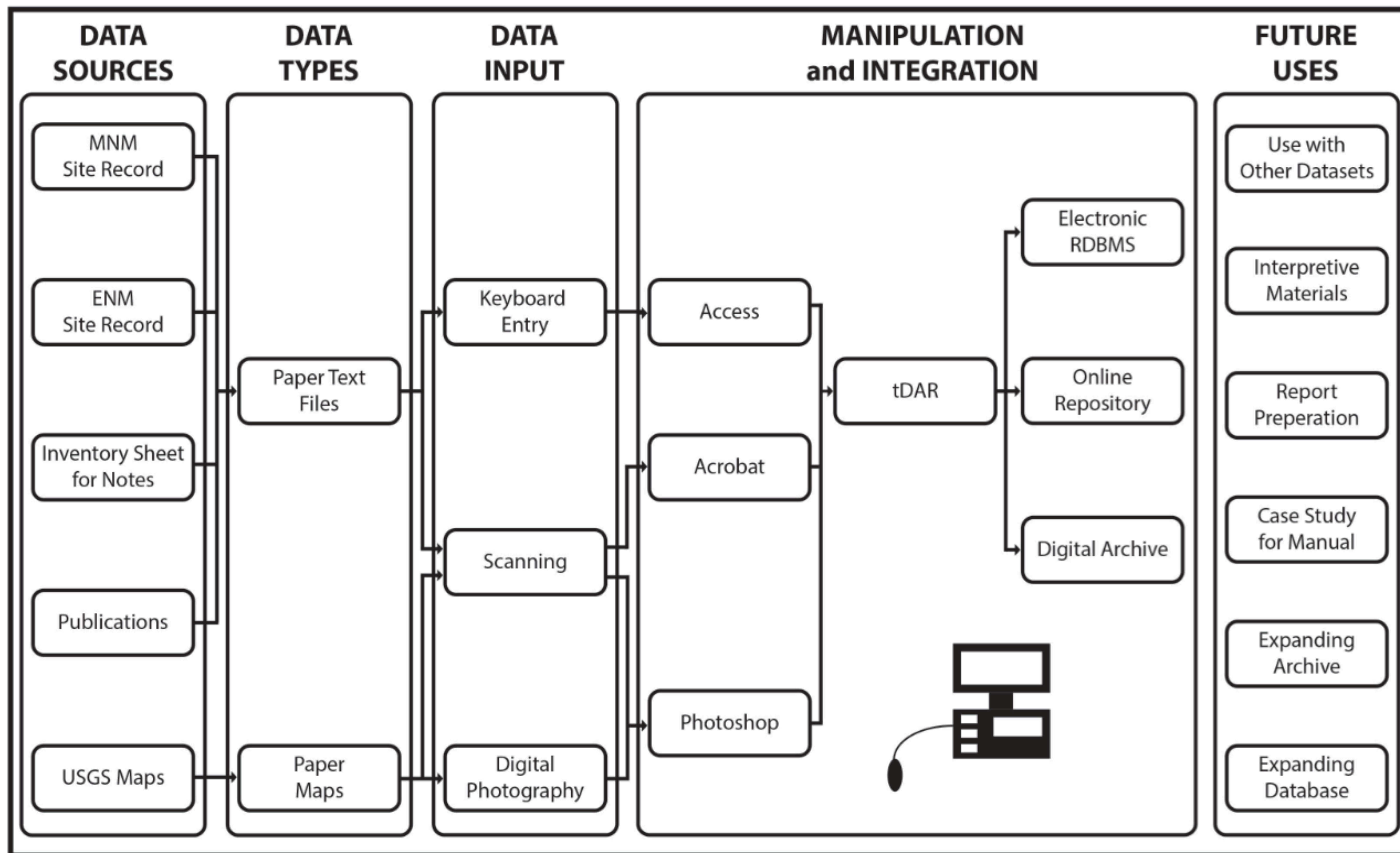
Additional research is necessary to continue locating and digitizing the various components of the AOP collection that are located across the United States. Currently, arrangements are being made to digitize and integrate the components of this collection at the National Anthropological Archives with those at ENMU into tDAR. Moreover, there are still components of the AOP legacy archaeological collection that are housed at ENMU that can be ingested into tDAR, such as the artifacts.

Additionally, this research brings forth the need for a professional niche that includes archaeology, archiving, and computer science – archaeological archiving. Archiving archaeological collections are becoming increasingly more difficult due the curation crisis, the proliferation of archaeological materials, and its cost to properly manage them⁵. Academics and professionals that have the skills of this niche can contribute immensely to this problem and help create big data in archaeology.

Bibliography

- ¹ Plaza, David M., "Digitizing the Anasazi Origins Project" Master's thesis, Eastern New Mexico University, 2013.
- ² Plaza, David M., "The Anasazi Origins Project Digital Archives Initiative: Transferring a Legacy Dataset to a Living Document Using tDAR." (2013) Report prepared for The Center for Digital Antiquity. Available via web <http://core.tdar.org/document/391353>.
- ³ McManamon, Francis P. and Kintigh, Kiehl W. "Digital Antiquity: Transforming Archaeological Data into Knowledge." *The SAA Archaeological Record* 10 (2010): 37-40.
- ⁴ Snijder, Chris, Uwe Matzaj, and Ulf-Dietrich Rieps. "Big Data": Big Gaps of Knowledge in the Field of Internet Science" *International Journal of Internet Science* 7 (2012): 1-5.
- ⁵ Childs, S. Terry and Eileen Corocan. "Managing Archaeological Collections: technical Assistance." (2000) Accessed January 21, 2013. www.cr.nps.gov/archaeology/collections/.

- Data Processing Flow Chart -



The Anasazi Origins Project Digital Archives Initiative mission is to aggregate all data associated with the AOP digitally onto platforms that is easily accessible and has the capability to disseminate. One of the main focuses is to virtually reunify and ingest the digital data of the AOP collection onto the Digital Archaeological Record (tDAR). This international online digital repository facilitates a living documents approach that allows data on this platform to be modernized and integrated with additional information from the various components of the AOP collection as well as other datasets. This project is in effort to begin the movement of the digitized data of the AOP into a central location with the goal to enhance its participation into the current archaeological dialogue.

This research was funded by the tDAR/DA Grant, ENMU Graduate Student Research Grant, and the ASNM Scholarship. Special thanks goes to John Montgomery, Rolando Marquez, David C. Batten, J. David Kilby, Bradford R. Hamann, and Josue Barrera of Eastern New Mexico University and Scott Thompson of Digital Antiquity for their assistances. This research poster is based on the tDAR/DA Grant report titled, "The Anasazi Origins Project Digital Archives Initiative: Transferring a Legacy Dataset to a Living Document Using tDAR."

