Andrew CharltonJones

LIS 545

1/28/24

*Data and Metadata Profile*

The data I will be discussing is a [collection of compiled observational data regarding European prehistoric archaeological digs containing jewelry and other adornments](https://zenodo.org/records/10545573). The data consists of four distinct datasets. These datasets are divided into two distinct categories of data: raw and processed. The raw datasets consist of two datasets, the first is sites: a dataset listing the sites where prehistoric adornment items were found. Then there is structures: a dataset listing the structure type and name where prehistoric adornment items were found. The processed datasets are: records: a dataset of adornment items found at these sites and structures records: a dataset displaying more in-depth information about the locations where these items were found. Despite similar sounding names all datasets contain unique data regarding the archaeological sites. All four datasets are available via download as CSV files. I have successfully downloaded/viewed them on Mac and PC. They are accessible via Numbers or Excel.

The data is from the [Prehistoric Europe’s Personal Adornment database](https://pepadb.us.es/app/). An EU-funded database for collecting and sharing European prehistoric site data. The data was “extracted from bibliographic resources”. The specific criteria for which archaeological resources were chosen and why is absent from both the Zenodo page and the PEPAdb. However, this omission may be rectified in the future as the database is far from complete.

This data has a slew of potential stakeholders. The first stakeholders that come to mind are archaeologists. Archaeologists can use this data to identify important sites for further excavation, relevant sites for their research, and previously investigated sites. Archaeologists are also the people creating and curating this data. Thus, Archaeologists have many incentives to ensure the data remains current and useful. The PEPAdb’s goals are to encourage this data to be shared and used by all. Therefore, scholars of all kinds are another stakeholder. Scholars in other fields could analyze and incorporate this data into their work and specialties. Given the PEPAdb’s goals, ensuring these stakeholders can use the data is a major goal. The EU is another stakeholder. PEPAdb is funded by the EU. The EU has an ongoing interest in ensuring this data is accessible and useful. An important concept to consider is that anyone with an internet connection is a potential stakeholder. Anyone can find and access this data. This is an important consideration especially when considering accessibility. This universality ensures that this data can/may be approached from a variety of perspectives. It also means that any culturally sensitive or identifiable information needs to be very carefully considered.

This data does not appear to have any use restrictions. The website and the Zenodo page don’t list any. There may be restrictions as there are funding documents on the Zenodo page in Spanish that I couldn’t read. However, it seems unlikely that these restrictions wouldn’t be mentioned elsewhere.

The metadata for the datasets is not located on the Zenodo page. However, it can be downloaded from the PEPAdb website. It is available in TTL form and requires specific software to open. The title metadata feature for the datasets is confusing due to their use of similar terms. However, the titles appear to be standardized making change unlikely. The dataset keywords only contain personal adornment omitting other potential access points like archaeology or prehistory. The metadata elements for the data vary by dataset. For example, sites only lists the site, whereas structures has elements like site, country, and structure type. While there are repeating metadata elements each dataset has different features. The metadata features are not explained or defined. Users have to learn what the metadata features represent via context clues, additional research, and/or expert knowledge. For example, adm1 and adm2 are not explained. It is through research one learns that these refer to two levels of administrative government. The metadata features are not very comprehensive. Even the most comprehensive datasets do not possess many or particularly in-depth features. Furthermore, all datasets feel like they could be expanded with additional metadata. Sites which only features one metadata feature could be expanded with additional features like state, climate, etc. The database is still a work in progress; perhaps more robust metadata features will be forthcoming.

The Zenodo page does not mention any metadata standards. However, the Prehistoric Europe’s Personal Adornment database states that the geographic metadata meets ISO metadata geographic standards. The raw and processed datasets meet RDF standards.

As addressed above the metadata feels both vague and lacking in explanation. Additional keywords regarding archaeology or prehistory should be added to the dataset metadata to increase findability. The metadata for the data would also be improved with additional metadata features. For example, features like culture associated with the sites, additional information regarding the digs, or descriptions of the adornments would be useful for researchers. Currently, this data paints a useful but broad picture. It tells us that 50 beads were found in a cave near Madrid. However, additional details like what shape those fifty beads were, is this site associated with a bead-making culture, and nearby sites with similar objects could be useful. This additional metadata would help users better understand these sites. I especially feel this additional metadata may be useful for users lacking familiarity with archaeology and prehistory. Currently, the lack of specificity and lack of explanation of terms may exclude some users from using the data to its full potential. The presence of additional explanatory metadata features may remove some barriers. It would also be helpful to enhance findability with additional keywords. For example, presently researchers looking for data on digs involving the beaker culture may not find this database.

Zenodo contains no citations of these datasets. However, the Zenodo page is recent, having been published on January 21, 2024. This data may simply be too new to have been cited yet. While I found many papers referencing the topic I found no papers that cited this data in the UW libraries. I advanced search using multiple terms and individually via subject, and keyword searches for adornment, prehistoric, archaeology, Prehistoric, Jewelry, and Prehistoric Europe’s Personal Adornment database. I also tried searching by author for the contributors listed on the Prehistoric Europe’s Personal Adornment database page. They had papers,but none citing this exact data.