

WLP DATA MANAGEMENT PLAN

DIGITAL DATA

The Shelby White and Leon Levy Program (WLP) for Archaeological Publications recognizes that publication may take many forms and follow different models. Digital data plays an increasingly important role in scholarship and the WLP encourages applicants to consider innovative modes of digital dissemination as a primary or secondary outcome of publication projects.

RATIONALE FOR DATA MANAGEMENT

Like the National Science Foundation (NSF) and the National Endowment for the Humanities (NEH), the WLP is committed to encouraging good practice with regard to the management, preservation, and access to research data. Most current publication projects involve the management of various forms of digital data (documents, data tables, digital images, GIS files, etc.). These data may represent invaluable and irreplaceable resources needing long term preservation. Similarly, these data resources may be valuable to the scholarly community for research and instructional purposes. Like NSF and NEH we ask grant seekers to plan for the preservation and access of data managed during their proposed publication projects.

ELEMENTS OF A DATA MANAGEMENT PLAN

1. Describe if and how you plan to disseminate digital products of your research (data tables, digital images, GIS files, etc.).
2. Describe if and how you intend to document and archive these digital resources for future reuse.
3. Describe data management work-flows to prevent data loss and to maintain data integrity and quality.

We encourage working with a "memory institution" such as a university library, disciplinary repository, dedicated data publisher, or similar organization dedicated to digital data dissemination and preservation. In order to permit legal data reuse and archiving, we recommend but do not require licensing of digital content using a [Creative Commons Attribution license](#). Researchers needing assistance with identifying a suitable archive and composing a data management plan should use this website: <http://dmp.cdlib.org/>.

DIGITAL DATA AND FIRST YEAR OUTCOMES

Efficient and effective data organization and documentation are key to a successful publication project. At the end of the first year, we ask grantees seeking renewal to demonstrate progress toward meeting publication goals. A well-organized corpus of digital content relating to the publication should be used as evidence for progress and sent to the WLP program for review. Review criteria will include:

1. Data tables need to have sufficient documentation to understand (fields need to be described, field values need to be decoded).
2. Descriptive vocabularies and terminologies need to be applied consistently (spelling, capitalization, plurals). Ideally, terms used in classification should be related to terminologies already in wide use or use by authoritative sources.
3. Image and other media files should not be stored within a database (in a binary or media field). Rather, these need to be stored as separate files in clearly and consistently organized directory structures. A data table needs to be provided that documents these media files and associates them clearly and unambiguously with identifiers to other entities in a dataset (archaeological contexts, artifacts, ecofacts, people, etc.).
4. Attribution information needs to be provided for all entities (images, observations / identifications, field documents).
5. To aid interoperability and help long term data preservation, researchers should have all data available ideally in open and non-proprietary file formats (or very widely used and understood proprietary formats). File formats should be used appropriately for the kind of content they represent. For example, tabular data should be in a tabular file format, and not expressed in an Acrobat PDF file.

Although we do not require digital publication and/or archiving of these digital resources, we do encourage dissemination and preservation of such digital research materials. This disposition of this content should be included in your project's Data Management Plan.

SAMPLE DATA MANAGEMENT PLANS

EXAMPLE 1: Project Using a National Data Repository

"As we work, we will make sure we regularly backup all digital data on multiple media. To improve data quality, we will use pick-lists and forms so data entry will be more consistent. We will also follow strict naming conventions on image files so their associations with object identifiers and context identifiers are clear.

Our project will archive data, including spreadsheets, database files, and image files with the Archaeology Data Service (ADS). The ADS is the leading digital repository of archaeological data, and is the national archaeological data repository for the United Kingdom. With a nearly 20-year track record, its expert staff closely adheres to best practices in digital data preservation. We will follow ADS recommendations for file formats and metadata (catalog and descriptive information) about the data we will submit to the ADS. Because the museum housing our project's collection imposes some intellectual property restrictions on

digital dissemination, we will release data using a Creative Commons Attribution – Non-Commercial license."

EXAMPLE 2: Project Using a Nonprofit-Managed Data Publishing Service

"My work will digitize approximately 2,000 pages of hand written excavation logs. I will store the digital image files in the TIFF file format which is non-proprietary and does not degrade image quality with compression. I will document these image files with an Excel spreadsheet.

The excavation diaries are housed in an archive that does not claim copyright ownership of them. Since they are 95 years old, they are suitable for release on the Internet as public domain resources. I will clearly indicate public domain status by using the Creative Commons Zero (CC-0) license. I will publish these digitized field-notes and their documentation with Open Context, an archaeological data publishing platform that archives data with the California Digital Library for long-term preservation."

EXAMPLE 3: Project Developing a Website and Using a University Digital Repository

"This publication project will develop a freely accessible website that will enable users to search through a variety of multimedia documentation. Our university IT support services will ensure that the website will remain operational for 5 years following the end of the grant period. For longer-term access, all the data and media files will be archived with our university's digital repository. The digital repository also provides persistent identifiers (DOIs) that ensure the content developed by our project can be cited securely in future scholarly works.

We will document the data created in this project using metadata ("information about information") standards developed by other leading digital data initiatives in archaeology, including the Archaeology Data Service, Digital Antiquity, and Open Context. All data and media files will be released immediately in open file formats under a Creative Commons Attribution license."