

Data Management Plan

This data management plan was created on September 3, 2013, for submission to the Office of Digital Humanities (ODH), National Endowment for the Humanities as required by ODH Guidelines in the interest of securing funding for this project. This is the first version of the data management plan associated with this data.

Types of Data

The proposed project is a workshop with an accompanying demonstration “sandbox” of potential technologies to be evaluated for a larger project on Afro-Caribbean labor, migration, and the Panama Canal. The following types of data will be produced: 1) a white paper outlining opportunities for further work; 2) a digital bibliography of sources related to Afro-Caribbean labor; 3) a website for the workshop; and 4) a machine image with the Open Source Indexing software installed, suitable to be run on the Amazon cloud computing infrastructure.

Data and Metadata Formats

The white paper will be maintained as a PDF file. The digital bibliography will be created and managed using the Zotero citation management software during the active life of the project. At the end of the project, the bibliography will be exported using the software’s built-in functionality as both a plain-text, UTF-8-encoded BibTeX file and as an RDF/XML file using terms from the Bibliontology suitable for migrating and preserving the bibliographical data. The workshop website will be built on a current version of the Wordpress content management software and will consist of PHP, HTML, and CSS files as well as images, videos, and other components. The machine image is a proprietary virtual appliance (a package format consisting of a file system, operating system and additional software) designed to be run on the Amazon cloud infrastructure.

Data Storage and Security

At least two copies of the data will be actively managed and stored during the life of the project. Data will be stored on password-protected servers managed by the University of Maryland College of Arts and Humanities. Data is backed up incrementally through a service provided by the Maryland Office of Information Technology, which has a proven record of and commitment to secure data archiving for the university. Backups occur nightly. In addition to backups hosted on university servers, data will be copied to tape and stored at a geographically-distant site through a service provided by Iron Mountain information services.

Access, Dissemination, and Preservation

The project white paper will be made available from the CHNA and MITH websites as soon as it is available and for at least one year afterward. The white paper will also be available from the University of Maryland digital repository in perpetuity. The digital bibliography will be made publicly available through the Zotero website. Archived copies will be available from the University of Maryland digital repository. MITH commits to hosting the project website for at least one year after the completion of the grant. After that time, a static version as preserved by the Archive-It tool will be preserved and made available through the Internet Archive as well as through the University of Maryland Libraries Digital Collections. Since the workshop website will

not contain significant interactive components, this static version should be sufficient for preservation of the intellectual content.

The machine image containing the OSI software will be made public on the Amazon cloud computing platform. This will allow other researchers to have access to a working version of this software. The proprietary nature of the Amazon machine image would present a risk to long-term preservation but as the software “sandbox” is only intended for evaluation and since the OSI tools are under active development, developing a long-term solution for project software—such as a machine image built on open source components for use on infrastructure provided by multiple different cloud hosting providers is not a priority for this project. MITH commits to providing the working machine image for the workshop and for as long afterwards as is practical given the technical constraints.

For the white paper, compiled bibliography, and website the project will use a license that supports cost-free re-use of the data without requiring specific clearance from the Project Director or other rights holders (e.g., Creative Commons Attribution).

Long-Term Preservation

Within three years from the end of the grant period, at least one copy of all data (except the proprietary machine image) will be permanently archived with the University Libraries at Maryland. Copies of the data may also be deposited with other suitable repositories as identified by the Project Director. No data will remain on servers controlled by MITH after three years from the end of the project. Data will remain publicly available through the libraries’ digital repository. The University of Maryland has an interest in investing in the management of data created by researchers affiliated with Maryland and in providing digital preservation services, such as file validation, integrity checks, and, if needed, format conversion.