Roles and Responsibilities

The data management plan will be implemented and managed by Stacy Reardon, under the project supervision of Bryan Wagner. Reardon will manage data and backup during the development phase and transfer data to institutional repository UC DASH for long-term permanent storage. Patty Frontiera and Susan Powell will be responsible for geospatial data. In the event that one of the responsible parties listed above is no longer involved in the project, responsibility for data management and the project website will fall to Principal Investigator, Bryan Wagner. In the event that Wagner is no longer involved in the project, responsibility for the website and project data will fall to the UC Berkeley D-Lab.

Expected Data

Data will consist of digitized primary sources, transcripts, translations, metadata, digital geographic data, the project website and any attendant custom code and design files, documentation for internal use, and tutorials for public consumption. In addition, the project will produce curriculum materials, conference materials, and scholarly publications.

Period of Data Retention

Processed primary source data will be made available immediately upon launch of the project website. Additional and newly acquired primary source data will be made available in a timely manner upon processing. Monographs or articles resulting from the project will be made available as soon as possible according to the embargo period of the publisher, and open access publishing will be preferred.

Data Formats and Dissemination

Materials for public consumption will be made available on the project website and archived in UC DASH. Image files will be made in TIFF format and stored in California Digital Library's Merritt repository for archiving, and copies will be made in JPEG for web display. Transcripts, translations will be stored in PDF form. Dublin Core metadata will be maintained in CSV and XML. Digital geographic data will be maintained in JSON files, with an eye towards making it extensible to other projects. Custom code will be made available on GitHub and archived in Merritt. Tutorials and curriculum materials will be disseminated in the most appropriate media and stored as PDF when possible; video content will be in MP4 format. All internal documentation will be stored in UC DASH as plain text and CSV files. Scholarly output will be archived in California Digital Library's open access repository whenever possible.

Data Storage and Preservation of Access

During the active project phase, all data will be backed up in UC Berkeley's Google Drive storage. In addition, GitHub will be used for any custom code development.

After the grant period, the UC Berkeley D-Lab will assume web hosting costs and essential maintenance for the project's online learning environment. The D-Lab will pay the campus fee for a Pantheon Pro Account ($900/year; web.berkeley.edu/web-hosting-pantheon). The D-Lab will also pay to have the project website regularly updated, maintained, and repaired. This commitment is calculated at five hours a month ($30/hour or $1800 annually).

A graduate student assistant will ensure that the website is crawled and archived through the Internet Archive at least monthly from the time it is launched through the duration of the project phase, and thereafter the PI will be responsible for periodical, regular web archiving.

In addition, all images, texts, metadata, documentation, and code components will be made available for public download through a free and open-source GitHub site as well as through the UC DASH Repository (dash.ucop.edu). UC DASH allows access via persistent URLs, offers tools for long-term data management, and permits permanent storage options. UC DASH has built-in contingencies for disaster recovery including redundancy and recovery plans.