

DATA MANAGEMENT PLAN

Expected Data

The data generated by this project will fall into six major types: documentation (procedure manual, ImageJ plugin documentation), captured and processed images, RTI images, annotation data, web pages, and software code. Email correspondence among the team will not be considered archival data. Additionally, instructional videos may be created.

Data Format

The procedure manual will be archived in PDF format. The documentation for the ImageJ Spectral RTI Toolkit plugin will be archived in HTML5. The captured and processed images will be archived as uncompressed 16-bit TIFF files with EXIF metadata describing contents, capture, and processing. PNG and JPEG compression may be used for web dissemination of images that will also be available without compression. RTI images will be archived in PTM (Polynomial Texture Map) and/or HSH (Hemispherical Harmonics) format. WebRTI, as a form of web dissemination, will use PNG and JPEG compression. Annotations will be stored in XML and HTML, with TEI standards for transcribed text. Shared Canvas manifests will use JSON. Image metadata will adhere to Dublin Core standards. Authority data will comply with MADS/RDF. The proposed Digital Public Library of America record will use OAI-ORE. Webpages will use HTML5 with CSS and Javascript. The ImageJ plugin will be written in Java. Developments to WebRTI will be in CSS and Javascript. Instructional videos will use the H.264, MPEG-4 AVC format at 1280x720 or 1920x1080 resolution (used by BluRay and web browsers).

Access to Data and Data Sharing Practices and Policies

The primary means of access to all data will be palimpsest.stmarytx.edu. This site will include the IIIF compliant image repository, navigable web pages for each manuscript page with image and annotations, the complete data archive (mostly TIFF images), documentation, and software code. Static URIs will follow the principles of Linked Ancient World Data. Code will also be available through GitHub. The white paper will also be available through neh.gov. Instructional videos will also be accessible on YouTube. The complete data archive will also be available on magnetic disk in the St. Mary's University Library Archives. The description of Spectral RTI will be submitted to Wikipedia. Other major discoverability and aggregation engines will be Google, WorldCat, Digital Public Library of America (conditional), and the Online Critical Pseudepigrapha. Access to the data will be completely unencumbered; no account registration or other impediments will be required.

Policies for Re-Use and Re-Distribution

The manuscript remains the property of the Biblioteca Ambrosiana and permission will be required for reuse in print media, but permission has been granted for all non-profit use (see Letter of Commitment). All other products can be used for any purpose with attribution (CC BY-SA).

Archiving of Data

The project server hosted by Amazon Web Services includes RAID redundancy and off-site backup. Additional copies will be kept on magnetic disk minimally by St. Mary's University Library Archives, the project director (Hanneken), and the Early Manuscripts Electronic Library (Phelps). Data integrity will be verified with MD5 checksums throughout the project. Syntax of filenames will be documented, as will correlation with IIIF identifiers.