

9. Data Management Plan

Roles and Responsibilities

This data management plan will be implemented and managed by Christine Ruotolo, in consultation with the Content Stewardship team at the University of Virginia Library. The University of Virginia Library will have long-term responsibility for the permanent retention of encoded texts and image files that comprise the bulk of the project data, as well as for preserving a static archival version of the project website.

Expected Data

The following data will be produced in the course of the project:

- TEI-encoded text files in XML format
- Representative page images associated with each text file, in JPG format with accompanying metadata
- A customized TEI schema and project-specific templates for text encoding
- A customized web interface, based on WordPress and incorporating the Anthologize plugin
- Technical guidelines and workflow documents

Period of Data Retention

The TEI-encoded text files created by the project and their associated metadata will be deposited in the University of Virginia Library's digital repository and will thus be added to the Library's permanent digital collections. The project website will be actively maintained for a minimum of five years beyond completion of the project. Workflow documentation and reports will be retained on GitHub for a minimum of five years beyond completion of the project.

Data Formats and Dissemination

The texts created for the project will be encoded in TEI and stored in XML format. They will be delivered online through a customized instance of WordPress, and will be made available to the public as they are completed. Descriptive metadata will be contained in the TEI header, which can easily be converted to other standard formats as needed. Representative page images created for the project will be stored as JPGs and described using Dublin Core or another widely recognized metadata standard.

Web hosting for the project website will be provided by the University of Virginia Library, which is currently investigating long-term hosting solutions for developing and preserving digital projects created by faculty.

Data Storage and Preservation of Access

Upon project completion, the WordPress site will be crawled using the Archive-It web archiving service, creating an archival snapshot of the finished project that will be added to the University of Virginia Library's Fedora-based digital repository for long-term preservation. The TEI-encoded text files will also be

ingested into the digital repository as stand-alone objects, and will be publicly accessible through the Library's discovery interface. The ultimate preservation destination for these materials is the Academic Preservation Trust (APTrust), of which the University of Virginia Library is a founding member.

Other data created by the project will be stored and disseminated via GitHub, a publicly accessible code repository. These data will include page images, schemas and markup templates, customized software, technical guidelines, and workflow reports.