

Data Management Plan

Data: All of the metadata on the ESSSS database is located on the Vanderbilt University Website, stored in a MySQL database which uses a Dublin Core based structure. This requires the production and storage of five versions of each image in order to produce the display. It is stored on network servers with multiple layers of protection and regular backups, and all files are copied onto virtual tapes. For additional security, the system is mirrored at an off-site location. In order to ensure the long-term preservation of these images and their metadata, however, we must encode them according to the Metadata Encoding and Transmission Standards (METS) in XML and then move them to the dark archive in the Digital Preservation Network (DPN), which allows indefinite storage in a failsafe repository. Only three versions of the images are created and stored for processing, allowing them to be optimized to require less storage and fewer resources. If funded, our staff will have sufficient support to encode all of the images and convert them for the DPN.

Interface: The current ESSSS website utilizes two distinct technological frameworks. The content management system is based on PHP (OmniUpdate) and the second uses a locally-developed PERL-based framework for presenting the actual images. Most other humanities databases and archives in use today have moved away from these individual frameworks and toward standardized, open-source systems using TEI or MySQL. This means that without significant manual labor, data from the ESSSS database cannot be effectively integrated with those of others. Despite this, Google Analytics demonstrates that the ESSSS database has a wide readership and that scholars from all over the world use this data in their research. In order to allow optimized data-sharing our staff will transfer the existing files with enriched and

standardized metadata to new and improved software platforms. ESSSS team members will transfer the two existing frameworks into a single system utilizing SobekCM open-source software (<http://sobekrepository.org>), which utilizes C# to process metadata stored as METS but also provides a crosswalk to TEI. Therefore, all of the information in our current system will be integrated into one software that is designed to work with the types of software the most current Digital Humanities databases utilize, making them interoperable.

Sharing: Every aspect of the database, from the metadata, to the images, interface, and the software used will be made freely available to anyone in the world. All source code developed in connection with this grant project will be released under the General Public License Version 3.0 (GPLv3) (<http://www.gnu.org/licenses/gpl-3.0.en.html>) and publicly distributed through Github. All metadata will be dedicated to the public domain via the Creative Commons Public Domain Declaration (CC0 1.0 Universal) (<https://creativecommons.org/publicdomain/zero/1.0/legalcode>). All images, both the high-resolution as well as the access images, will likewise be dedicated to the public domain via CC0 1.0 Universal (unless prior licensing restrictions have been attached). Metadata will be made available through Github with direct access to images.