

Data Management Plan

Since 2008 the website, www.slavevoyages.org that this project is designed to sustain and enhance, has provided free and unrestricted access to three databases that have become standard sources for anyone interested in the transatlantic slave trade. The first is a set of nearly 35,000 transatlantic slave voyages (termed “TSTD”) containing 99 variables, the second, a database of 11,709 annual/regional estimates of the slave trade (termed “Estimates”) and the third provides the personal details of 92,000 Africans rescued from slave vessels in the last half century of traffic (termed “Names”). Not only are these data viewable, but they may be searched, selected and analyzed, and any part of the data or the output of subsequent analysis may be downloaded into csv or Excel files. Given that the transatlantic slave trade ended in 1867 there are no legal or ethical restrictions on the use to which these data may be put and there are no retention requirements. TSTD and Estimates contain imputed variables as well as data variables. In each case the website offers full explanations as to how the imputed variables were generated and these explanations, too, may be downloaded without restriction. The downloads page at <http://www.slavevoyages.org/tast/database/download.faces> provides versions of the three databases in a variety of formats. The refurbished website will continue these data management policies.

This project will result in the creation of an additional database of slave voyages. These sailed between one port in the Americas to another and date from the same era as TSTD. It will be termed the Intra-American slave trade database. Its structure will be identical to TSTD in that it will have a mix of data and imputed variables and the algorithms for calculating the latter are also the same. By 2007 we estimate that it will comprise at least 10,000 records and contain 99 variables. The project team will spend seven months preparing the dataset and will make it available for use on its own interface that will be built as the database itself is created. Once the new site is launched then it will take its place with the other three data sets and be subject to the same unrestricted public access. A copy will be made available on the website’s download page in Excel, SPSS and csv. Also available on this page is the SPSS codebook. Of the 99 variables in the new database, 32 are coded.

The website will be stored and maintained on servers managed by the Emory University Library & Information Technology Services (LITS). The re-engineering of the site will include provisions to enable effective web crawling for archiving purposes, which will significantly improve on the current site. As of now, the site cannot be effectively archived via web crawlers due decisions made about Java implementation. The re-coded site will address these web archiving hindrances, allowing the website to be included in regularly scheduled crawling and archiving of university domains, and archived in the Emory University Archives Web Archives. The datasets and all accompanying documentation (codebooks, methodology and guides from the website) will be archived in the Emory Dataverse, which participates in the Dataverse Network, an open source framework for archiving and disseminating data developed at Harvard University. The Dataverse Network utilizes a secure and geographically distributed digital preservation system using the LOCKSS software created by Stanford University and under the guidance of the Data-PASS project. Copies of the site and the data it draws on are also located at the Wilberforce Institute for the Study of Slavery and Emancipation, Hull University, UK, (under the guidance of Richardson), at the Fundação Casa de Rio Barbosa of Rio de Janeiro, and at Victoria University of Wellington (under the guidance of Behrendt). Emory LITS is negotiating participation in an established national digital preservation network, which would provide a long-term multi-site dark copy of the data and website content, in the event of total failure of the university’s systems.