

## 7. Data Management Plan

### Data to be Generated

Type of Data	When shared?	Under what conditions?
Open Source computer code associated with tool, including algorithms and interface.	At conclusion of the start-up project, when initial testing has been completed.	Code will be freely available.
Dataset from ProQuest “Black Abolitionist Papers” database, to be used as test data.	ProQuest materials; cannot be shared. See next item.	N/A.
Dataset of freely-available abolitionist newspapers, to be released with computer code.	At conclusion of the start-up project, when initial testing has been completed.	Data will be freely available; see “Appendix D” for title list and sources.
Multimedia progress reports.	At the time of their writing, throughout the duration of the project.	The progress reports will be freely available on the Digital Humanities Lab website.
White paper.	After the project has been completed.	The white paper will be freely available on the Digital Humanities Lab website.
Multimedia report posted on the Digital Humanities Lab blog.	After the project has been completed.	The multimedia report will be freely available on the Digital Humanities Lab website.
Final report to NEH.	At the conclusion of the project.	Dissemination of the final report will be the responsibility of the NEH.

### Period of Data Retention

Data will be retained for 5 years beyond the completion of the start-up phase of TOME. Formal reports will be publically available within 1 year of project completion, on the Digital Humanities Lab website and on SMARTech (<http://smartech.gatech.edu/>), the Georgia Tech DSpace repository. Copies of the data will also be stored long-term on SMARTech.

### Data Formats and Dissemination

Computer code (Java and JavaScript) and public dataset (plain text) will be available as open source on SMARTech and GitHub, a publically-accessible code repository. All metadata associated with test dataset files will also be made available through these venues. Reports will be made available in PDF format and disseminated via the Digital Humanities Lab website (<http://dhlabs.lmc.gatech.edu>).

### Data Management and Maintenance

All computer code and public test data will be stored in Github, where Eisenstein’s previous projects have been stored. All other data, including test data that cannot be shared, as well as any reports and publications, will be stored in SMARTech, Georgia Tech’s online institutional repository for faculty and researchers.

SMARTech is part of the MetaArchive Cooperative distributed preservation network, a large-scale effort for the preservation of electronic scholarly materials through the Library of Congress’s National Digital Information Infrastructure and Preservation Program (NDIIPP). For the details of its preservation methods, see: <http://smartech.gatech.edu/policy>.