

## Data Management Plan

### Data to be Generated

Type of Data	When Shared?	Under What Conditions?
Open source computer code for user interface and functionalities of the tool.	After completion of initial start up phase.	Code will be freely available on the project website and Github.
User generated texts, comments, and forum-type discussions.	This data will be available immediately after input.	Users will specify whether their content is public or accessible only to a select group of users within the network.
Metadata from user generated texts, comments and networks.	After launch of prototype.	Automated text mining functionalities will publish trending topics, threads and papers while maintaining user privacy.
User assessments and team evaluation.	After launch of prototype; ongoing basis as data is collected.	Both survey-based and focus group evaluations will be freely available on the project website.
White Paper.	After the conclusion of the initial grant funded period.	The White Paper will be freely available on the project website in PDF format.
User manual and tutorials.	After the launch of the prototype.	User manuals and tutorials will be freely available on the project website and with the redistributable software package.
Final Report to NEH.	At the conclusion of the initial grant funded period.	Dissemination of the final report will be the responsibility of the NEH.

**Period of Data Retention** -- All relevant data will be deposited continuously into the CUNY Academic Commons ("Commons") Server for long term storage. Formal reports, tutorials and software code will also be freely available on the project website for a period no less than five years. User-generated content will be maintained on the project website according to the privacy and retention preferences set by individual users, but will be accessible for a period of no less than five years.

**Data Formats and Dissemination** -- Software code will be made available for access and modification under an open source license on Github, a publicly accessible code repository, as well as on the project website. Formal reports will be made publicly available in PDF format on the project website at the end of project completion. All metadata associated with user generated content will be available, while maintaining user privacy, through the project website. Users will determine the extent which their user-generated content and associated metadata is publicly available. In addition, users will specify and maintain the copyright and permissions associated with their individually contributed content, and thus, all user-generated content will be easily extractable by the user should they wish to remove their content from the project website. Tutorials to assist users with privacy, permissions, and data extraction will be freely available on the project website at project launch.

**Data Management and Maintenance** -- All computer code, including versioned records of all changes will be licensed under the GNU General Public License and stored in Github, where existing code for Commons In A Box, the underlying software infrastructure for the Social Paper, is stored. All other data, such as user generated content, official reports and tutorials, will be hosted on the Commons server, which is administered and housed on site by the Information Technology (IT) Department at the CUNY Graduate Center (GC). The IT Department creates a data backup every day at 6am by exporting all databases and all files on the Commons server to a co-located backup server at the GC. Each daily backup is kept for 28 days. Additionally, the Graduate Center has an identical fail-over physical server located at Baruch College, to which it synchronizes data daily at 5am. The Domain Name Service (DNS) record for the Commons, commons.gc.cuny.edu, is hosted through the cloud-based Amazon Route 53 DNS service, so the Commons can be switched over to the Baruch College servers should GC facilities become unavailable.