

---

## Plan Overview

*A Data Management Plan created using DMPTool*

**Title:** Multimedia Text Annotation for Students

**Creator:** DMP dmpcurator

**Affiliation:** University of California, Office of the President (UCOP)

**Funder:** National Endowment for the Humanities (neh.gov)

**Funding opportunity number:** 8328

**Template:** NEH-ODH: Office of Digital Humanities

**Last modified:** 05-28-2014

### Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# Multimedia Text Annotation for Students

---

## Roles and responsibilities

The Data Management Plan should outline the rights and obligations of all parties with respect to their roles and responsibilities in the management and retention of research data. It should also consider changes to roles and responsibilities that will occur should a project director or co-project director leave the institution or project.

## Expected data

The Data Management Plan should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

Type of Data: Open Source computer code associated with tool, interface, and server-side components development. When Shared: At conclusion of the start-up project, when initial testing has been completed. Under What Conditions: Code will be freely available. Type of Data: User-generated texts, annotations, and electronic multi-media collections generated during the testing phase. When Shared: None of these data will be publically available until the conclusion of the start-up project. Under What Conditions: These data will be made available at the discretion of the creator, who will control access via privacy settings. Copyrighted data will be dynamically excluded from collections and documents made available to the public. Type of Data: Assessment data generated during the testing phase. When Shared: Aggregated data will be shared via the white paper and final report to NEH. Under What Condition: No information will be shared that could identify individuals participating in the assessment process. Type of Data: White Paper When Shared: After the project has been completed. Under what conditions: The white paper will be freely available to the public via the project website. Type of data: A multimedia report posted on the project website. When shared: After the project has been completed. Under what conditions: The multimedia report will be freely available on the project website. Type of Data: A final report to the NEH. When shared: At the conclusion of the project. Under what conditions: Dissemination of the final report will be the responsibility of NEH.

## Period of data retention

NEH is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by discipline. It is strongly committed, however, to the underlying principle of timely access. In their DMP applicants should address how timely access will be assured.

Data will be retained for 5 years beyond the completion of the start-up phase of Annotation Studio Reports and aggregated data will be publicly available within 1 year of project completion, via the project website, copies of the data will be stored long-term in DSpace@MIT. <http://dspace.mit.edu/>

## Data formats and dissemination

The Data Management Plan should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

Computer code will be available as open source in a publicly accessible code repository (GitHub). Reports will be made available in PDF format and disseminated via the project website and through DSpace@MIT. All metadata associated with media documents in shared multimedia collections in Annotation Studio will be freely available on the Annotation Studio website, only copyright cleared media and text documents will be accessible to the public

## Data storage and preservation of access

The Data Management Plan should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. These can include third party facilities and repositories."

All computer code will be stored in Github, where existing code base for Metamedia (Annotation Studio's precursor) and other HyperStudio projects and tools has been stored. <https://github.com/>. All other data, including user generated texts and annotations, (exportable as TEI/XML files), assessment data, reports, and publications will be stored in DSpace, MIT's online institutional repository for faculty and researchers. Copyright-cleared media documents will be stored in DSpace along with their respective collection and student interaction information. DSpace@MIT identifies two levels of digital preservation: bit preservation, and functional preservation. Bit preservation ensures that a file remains exactly the same over time – not a single bit is changed – while the physical media evolve around it. Functional preservation goes further: the file does change over time so that the material continues to be immediately usable in the same way it was originally while the digital formats (and the physical media) evolve over time. DSpace@MIT insures permanent data preservation in a secure and searchable archive, and is managed by MIT Libraries.