
Plan Overview

A Data Management Plan created using DMPTool

Title: A Community Workbench for Studying Enlightenment Media

Creator: Molly Bostrom

Affiliation: University of Minnesota (twin-cities.umn.edu)

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

Funding opportunity number: 20210115-HAA

Template: NEH-ODH: Office of Digital Humanities

Project abstract:

We seek funding under the criterion of "creating or enhancing experimental, computationally-based methods, techniques, or infrastructure that contribute to the humanities." With a Level III grant, we will build upon our Level II activities focused on developing new understandings of the Cérémonies et coutumes religieuses de tous les peuples du monde (Religious Ceremonies and Customs of All the Peoples of the World, hereafter CCR)—our humanities subject—and its place within Enlightenment studies by expanding our robust, multidisciplinary, and international community of academic scholars, educators, librarians, graduate students, and book collectors. Advancing that research and community building requires the development of a digital humanities tool—Mirador Workbench—with capabilities currently unavailable to not only CCR researchers but the broader digital humanities community.

Start date: 09-01-2021

End date: 08-31-2024

Last modified: 03-15-2021

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

A Community Workbench for Studying Enlightenment Media

Roles and responsibilities

The DMP should clearly articulate how sharing of primary data is to be implemented. It 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 if a project director or co-project director leaves the institution or project. Any costs stemming from the management of data should be explained in the budget notes.

Both the University of Minnesota and Stanford University will produce data that requires management and preservation. Minnesota will manage and preserve data related to the *CCR Portal* and *N&RCCR*. Stanford will manage and preserve data related to Mirador Workbench.

Expected data

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

Project directors should address matters such as these in the DMP:

- the types of data that their project might generate and eventually share with others, and under what conditions;
- how data will be managed and maintained until shared with others;
- factors that might impinge on their ability to manage data, for example, legal and ethical restrictions on access to non-aggregated data;
- the lowest level of aggregated data that project directors might share with others in the scholarly or scientific community, given that community's norms on data;
- the mechanism for sharing data and/or making it accessible to others; and
- other types of information that should be maintained and shared regarding data, for example, the way it was generated, analytical and procedural information, and the metadata.

University of Minnesota

- A [International Image Interoperability manifest](#) list of all known copies of the *CCR*
- Journal articles and their related metadata published in *N&RCCR*

Stanford University

Mirador Workbench software and documentation: The software portion shall consist of software application code documentation, developed in publicly-accessible GitHub repositories during the project and following its conclusion. The software supports the creation and maintenance of the advantages used by the *CCR Portal* and *N&RCCR*.

University-hosted instance of Mirador Workbench service: Data stored within and produced by the Stanford-hosted instance of Mirador Workbench will include the following: configuration of Mirador workspaces and “vantages”; information about externally-hosted IIF resources in user libraries, workspaces, and projects; and annotations on IIF resources. Hosted data will include the instances of Mirador used for the *CCR Portal* and journal. Data produced within the service will be shareable either through published “vantages” of a user-configured Mirador Workbench workspace (for linking from or embedding within a website), or through user-initiated exports (for loading within another instance of Mirador Workbench). The data within the Stanford-hosted instance will be hosted on multiple virtual machines behind a load balancer, with an associated database managed as a separate virtual machine. Application data and associated databases receive nightly incremental and weekly full backups, retained for up to one month on a separate local storage environment. Offsite replicas are stored in IBM Spectrum Protect disk storage in a Texas data center.

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.

Stanford University

Mirador Workbench software and documentation: Stanford Libraries commits to develop the software openly during the project and to continue to make the software and documentation available after the conclusion of the project. Software dependencies will be updated through regular maintenance periods.

University-hosted instance of Mirador Workbench service: Stanford thus will commit to supporting and hosting the instance of Mirador Workbench, including security updates and maintaining the service on Stanford infrastructure, for at least three years following the project's conclusion. Stanford expects to continue new feature development and bug fixes for the Mirador Workbench service for at least two years following the project's conclusion. Stanford Libraries and its staff are active participants in the IIF Consortium and community and are invested in its success. However, Stanford Libraries retains the right to transition the hosted instance of Mirador Workbench into a “read-only” mode after three years following the period of performance.

Data formats and dissemination

The DMP 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.

University of Minnesota

All volumes of this 18th-century text are in the public domain. Since we will be pulling in IIF-compliant copies of the *CCR* using the IIF manifest rather than attempting to host them ourselves, ownership and rights determinations will remain with the host institutions. Our site will merely provide a view of the content; any users seeking to reuse the work would have to contact the host institution for permission. As is standard with IIF protocols, rights information, if available, would be put in with the rest of the metadata and would be visible to the user in the metadata view. The manifest list will be formatted and preserved as a single .csv file and the journal articles and metadata will be formatted and preserved as .pdf, XML, and HTML packaged separately for each article.

Stanford University

Mirador Workbench software and documentation: **The software and documentation shall consist of application code written in Ruby and JavaScript, and documentation written in Markdown and/or HTML. All software products released under the Apache Software Foundation License, version 2.0, with copyright ownership held by Stanford University. All documentation will be released under a CC BY 4.0 license through the project's GitHub repositories.**

University-hosted instance of Mirador Workbench service: Data produced by the Stanford-hosted instance of Mirador Workbench will be managed within the instance's data stores (e.g. a relational database and associated Solr server). Annotations created within, and IIIF resources used within Mirador Workbench will be serialized and exportable to IIIF-compliant JSON-LD. Other data from the application will be serialized and exportable as JSON and will be suitable for import into another instance of Mirador Workbench. In addition, public workspaces will be shareable and embeddable as web pages. As a virtual research environment, data produced within the Stanford-hosted instance will only be published when data owners choose to do so. The Stanford team will ensure the instance complies with information security standards, such as the [University IT Risk Classification](#).

Data storage and preservation of access

The DMP should describe physical and cyber resources and facilities that will be used to effectively preserve and store research data. These can include third-party facilities and repositories.

University of Minnesota

The IIIF manifest list is a list of uniform resource identifiers that allow the CCR Portal (Currently hosted in Mirador 3.0, eventually to be hosted by Mirador Workbench) to display digital copies of the CCR. It will be maintained in Google Drive and backed up on local drives as it is gathered. A .csv copy of the list at the end of the period of performance will be preserved and made accessible in the [Data Repository for the University of Minnesota](#) (see the [DRUM preservation policy](#)) *N&RCCR* will be hosted by Public Knowledge Project and their installation of [Open Journal Systems 3.2.1](#). *N&RCCR* will only include publication data and the associated metadata. Metadata will be created and submitted by contributors (authors and journal editors). Published file formats will include EPUB, HTML, and PDF and will be available to readers on Manifold, [hosted by Cast Iron Coding](#). Publication metadata will be submitted to Crossref via Crossref's 4.4.2 metadata deposit schema in order to assign published articles a Digital Object Identifier. XML deposits will be facilitated through Open Journal Systems' Crossref XML Plugin. Published articles will be preserved in [Portico](#) under an agreement with Minnesota (see the [Portico preservation policy](#)) with Libraries Publishing committed to article access for *N&RCCR* indefinitely from the signing of the title's MOA forward. The *CCR Portal* will point to original existing digital versions of the text; therefore, preservation of the end points will be maintained by the host institutions. The majority of these host institutions are galleries, archives, libraries, or museums with a shepherding responsibility for the digital content.

Currently, the CCR Portal is hosted by the University of Minnesota Libraries on a stock version of Mirador 3.0. In year two of the period performance, the *CCR Portal* will transition from Mirador 3.0 to Mirador Workbench. The underlying data (i.e., the IIIF manifest) will be preserved in perpetuity in the Data Repository for the University of Minnesota and the underlying code for Mirador 3.0 will remain accessible via Stanford, but the portal itself will transition from Minnesota to Stanford and the existing portal will retire.

Stanford University

Mirador Workbench software and documentation: All software and documentation will be managed in openly accessible GitHub repositories and will be accessible without authentication. Development of the software and documentation will continue in the existing repositories. Upon project completion, archival release versions will be stored in a repository (either Stanford Digital Repository, or a software-focused repository platform such as Zenodo).