Data Management Plan

Letters of the Revolution Project

Contact person: **Thomas Hoekstra** (*tthomas2310@hotmail.com,*
0000-0002-4114-8373)
University of Groningen

Based on: *Common DSW Knowledge Model, 2.4.0 (dsw:root:2.4.0)*

Project phase: *After Finishing the Project*

Created by: **Thomas Hoekstra** (*tthomas2310@hotmail.com*)

Generated on: *11 Nov 2022*

Data Management Plan created in Data Stewardship Wizard «ds-wizard.org»

Projects

We will be working on the following projects and for those are the data and work described in this DMP.

The Letters of the Revolution Project: Correspondence Networks of the U.S. Founding Fathers

Start date: *2022-02-08*

End date: *2022-11-18*

Funding: *Did not apply for any funding yet.*

Section A: Data Collection

1. What data will you collect or create?

Data formats and types

We will be using the following data formats and types:

- **Comma-separated Values**

  It is a standardized format. This is a suitable format for long-term archiving. We expect to have 2.5 GB of data in this format.

- **Gephi**

  It is a standardized format. We are aware that this is not a suitable format for long-term archiving. We will have only a small amount of data stored in this format.

- **Scalable Vector Graphics**

  It is a standardized format. This is a suitable format for long-term archiving. We will have only a small amount of data stored in this format.

- **Portable Document Format**

  It is a standardized format. This is a suitable format for long-term archiving. We will have only a small amount of data stored in this format.

2. How will the data be collected or created?

There will be no instrument dataset in this project.

Data storage and file conventions

We will use a filesystem with files and folders with the following folder conventions:

- There will be a **folder for each sample/subject**.

We will not be storing data in an "object store" system.

We will not use a relational database system to store project data.

We will not use a graph database for data in the project.

We will not be storing data in a triple store.

Section B: Documentation and Meta-data

3. What documentation and meta-data will accompany the data?

List of data to be published is given in Section E, Question 9. This also includes information about catalogs where the data can be found. Information about data types used is given in Section A, Question 1.

We will use other solution than (electronic) lab notebooks to make sure that there is good provenance of the data analysis: All steps are described in Thesis and README file on Gephi.

Section C: Ethics and Legal Compliance

4. How will you manage any ethical issues?

5. How will you manage copyright and Intellectual Property Rights (IPR) issues?

We will be working with the philosophy *as open as possible* for our data.

All of our data can become completely open immediately.

Limited embargo will not be used as all data will be opened immediately.

Our data is legally not copyrightable, there is no legal owner.

For the reference and non-reference data sets that we reuse, conditions are as follows:

- **Founders Online** – freely available for any use (public domain or CC0).
- **People of the Founding Era** – freely available for any use (public domain or CC0).

Section D: Storage and Backup

6. How will the data be stored and backed up during the research?

Data that project members themselves store adequately backed up and traceable. Therefore data are protected against both equipment failure and human error.

7. How will you manage access and security?

Project members will not store data or software on computers in the lab or external hard drives connected to those computers.They will not carry data with them (e.g. on laptops, USB sticks, or other external media). All project web services are addressed via secure HTTP (https://...).

The possible impact to the project or organization if information is lost is small. The possible impact to the project or organization if information is leaked is small. The possible impact to the project or organization if information is vandalised is small.

Section E: Selection and Preservation

8. Which data are of long-term value and should be retained, shared, and/or preserved?

We plan to produce the following datasets:

- **Letters of the Revolution Project Dataset** (published) – This data set will be kept available as long as technically possible.

9. What is the longterm preservation plan for the dataset?

- **Letters of the Revolution Project Dataset** (published)
  The distributions will be stored in:
  - Our institutional repository.
  - Domain-specific repository: [GitHub](). We don't need to contact the repository because it is a routine for us.

None of the used repositories charge for their services.

Section F: Data Sharing

10. How will you share the data?

- **Letters of the Revolution Project Dataset**
  The dataset has the following identifiers:
    - URL: **https://github.com/Rohrym/Letters_of_the_revolution_project**

  The distributions will be available as follows:
    - Open (shared with anyone) using our institutional repository. The distribution will be available under the following license:
        - Freely available for any use (public domain or CC0).
    - Open (shared with anyone) using a domain-specific repository: GitHub. The distribution will be available under the following license:
        - Freely available for any use (public domain or CC0).

Information about used repositories (i.e. where will potential users find out about the data) is provided in Section E, Question 9.

Embargo on the data is described in Section C, Question 5, and Section F, Question 11.

11. Are any restrictions on data sharing required?

Ethical and legal restrictions are documented under Section C. We have used the Data Stewardship Wizard, which made us aware of options to minimize the restrictions.

No data sharing agreement will be required.

Section G: Responsibilities and Resources

12. Who will be responsible for data management?

Thomas Hoekstra is responsible for implementing the DMP, and ensuring it is reviewed and revised.

Thomas Hoekstra is responsible for reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use and maintenance within a data centre or repository.

Thomas Hoekstra is responsible for finding, gathering, and collecting data.

Thomas Hoekstra is responsible for maintaining the finished resource.

13. What resources will you require to deliver your plan?

To execute the DMP, no additional specialist expertise is required.

Charges applied by data repositories (if any) are mentioned already in Section E, Question 9.