Plan Overview

A Data Management Plan created using DMPonline

Title: Scan of Open Research Data at Utrecht University (SWORDS@UU)

Creator: Kavan Quach

Principal Investigator: Christopher Slewe, Keven Quach

Data Manager: Jonathan de Bruin, Christopher Slewe, Keven Quach

Project Administrator: Anna-Lena Lamprecht, Jonathan de Bruin

Affiliation: Utrecht University

Template: Utrecht University DMP

Project abstract:

SWORDS@UU aims to collect data and software published by UU researchers and analyze the output on FAIRness.

ID: 75025

Start date: 15-03-2021

End date: 30-06-2021

Last modified: 31-03-2021

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 customise 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

Scan of Open Research Data at Utrecht University (SWORDS@UU)

Data Collection

1.1 Will you re-use existing data ?

If yes: explain which existing data you will re-use and under which terms of use.

- No, I will be collecting/generating new data

1.2 Describe your data.

Fill the table below with a brief description of the data, including the type, format and volume.

The data description can be seen in the Research plan in detail.

Data Documentation

2.1 Describe the documentation and metadata that you will use to to make your data reproducible and interoperable.

Describe which files you will provide, along with a brief description of the information they will contain, to make your data reproducible and interoperable. Describe the information that you will provide to make the data items in questions 2.1 reusable and interoperable. If using a specific metadata standard, please mention this below.

Provided files:

- GitHub repository URLs
- Generated data of the repositories

- Analysis results
- Code for data generation
- Code for analysis

Based on an initial file of repository URLs, data will be generated. The code to reproduce the data generation will be made publicly available such that everyone can follow the steps. The same procedure will be taken for the analysis code.

2.2 Describe the folder structure you will provide to make your data reproducible and interoperable.

Describe the folder structure, naming conventions and/or version control you will use for this project.

TODO: This will be filled out while the project is ongoing.

Data Storage

3.1. Select the storage solution where you will store and back-up your data.

Select the locations where your data will be stored. You may select more than one. Please describe the storage solution and the backup strategy of your storage solution if it does not appear in the list below.

- Other (please specify below)

A snapshop of the data will be published to GitHub. Upon finalization of the project, data will be additionally stored on Zenodo.

Data Privacy and Security

4.1 Will you be collecting or using personal data ?

Personal data is any data which, alone or in combination with other information, can identify a living person. Such data must abide by the GDPR and requires additional safeguards and documentation to be processed lawfully.

- No, I will not collect and/or use personal data

4.2 How will ownership and intellectual property rights of the data be managed?

Describe who controls access to the data and who determines what is done to the data.

The data will be made publicly available.

Data Selection, Preservation & Sharing

5.1 Describe the data you will be preserving and the storage solution where it will be preserved?

Describe which data will be preserved under long-term storage. You may refer back to the data described in question 1.2 to specify which data will be preserved. Explain where you will preserve your data, and how procedures are applied to ensure the survival of the data for the long term.

All collected data will be stored on GitHub and Zenodo. The data will be kept indefinitely.

5.2 Describe the data you will be sharing and the repository where it will be shared?

Describe which data you will be sharing. Select where you will make your data findable and available to others. If selecting "Other" please specify below which repository and provide a URL.

Please also write below if you will apply any conditions to the re-use of your data. (i.e. Creative commons license or Data Transfer Agreement).

- Zenodo

5.3 Are specialized, uncommon or expensive software, tools or facilities required to use the data?

Please list any specialized, uncommon or expensive software, tools or facilities that are absolutely required to obtain, use or handle your data, if any.

No, all the data can be accessed by free, open-source or non-proprietary software. Data will be stored in a CSV format. Python will be used for data generation and analysis.

Data Management Costs and Resources

6.1 What are the foreseeable research data management costs and how do you expect to cover them ?

Please specify the known and expected costs involved in managing, storing and sharing your data. Also explain how you plan to cover these costs.

There are no foreseen data management costs.

6.2 Who will be responsible for data management?

Please specify who is responsible for updating the DMP and ensuring it is being followed accordingly.

The students Christopher Slewe, as well as Keven Quach will be responsible for maintaining the DMP.

6.3 State if you contacted an RDM consultant from Utrecht University to help you fill out your DMP.

Please list their name and date of contact.
This is mandatory for NWO grants.

No RDM consultant has been contacted.