Plan Overview

A Data Management Plan created using DMPonline

Title: Coding in public: recognising vulnerability as a barrier to reproducibility

Creator: Jools Kasmire

Principal Investigator: Dr. J. Kasmire

Data Manager: Dr. J. Kasmire

Contributor: Louise Capener, Nadia Kennar

Affiliation: University of Manchester

Funder: Economic and Social Research Council (ESRC)

Template: ESRC Template Customised By: University of Manchester

ORCID iD: 0000-0003-2684-6330

ID: 142314

Start date: 01-01-2024

End date: 28-06-2024

Last modified: 19-01-2024

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

Coding in public: recognising vulnerability as a barrier to reproducibility

Manchester Data Management Outline

1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- None of the above

2. Is The University of Manchester collaborating with other institutions on this project?

- Yes - Part of a collaboration and owning or handling data

The researchers are members of both the University of Manchester and the UK Data Service and so these two organisations can be understood to be collaborating.

3. What data will you use in this project (please select all that apply)?

- Acquire new data

4. Where will the data be stored and backed-up during the project lifetime?

- Other storage system (please list below)

All data will be stored in University of Manchester sharepoint sites. Data suitable for sharing will be stored in a GitHub repository within the UK Data Service's GitHub account.

5. If you will be using Research Data Storage, how much storage will you require?

- Not applicable

6. Are you going to be receiving data from, or sharing data with an external third party?

- Yes

As part of the research project, participants will be made aware of the UK Data Service's GitHub account and will be told that this (and all other reproducible reserach projects that we run) can see the appropriate for sharing data on those repositories.

7. How long do you intend to keep your data for after the end of your project (in years)?

- 5 - 10 years

Guidance for questions 8 to 13

Highly restricted information defined in the Information security classification, ownership and secure information handling SOP is information that requires enhanced security as unauthorised disclosure could cause significant harm to individuals or to the University and its ambitions in respect of its purpose, vision and values. This could be: information that is subject to export controls; valuable intellectual property; security sensitive material or research in key industrial fields at particular risk of being targeted by foreign states. See more examples of highly restricted information .

If you are using 'Very Sensitive' information as defined by the Information Security Classification, Ownerships and Secure Information Handling SOP, please consult the Information Governance Office for guidance.

Personal information, also known as personal data, relates to identifiable living individuals. Personal data is classed as special category personal data if it includes any of the following types of information about an identifiable living individual: racial or ethnic origin; political opinions; religious or similar philosophical beliefs; trade union membership; genetic data; biometric data; health data; sexual life; sexual orientation.

Please note that in line with data protection law (the UK General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as practically possible. You must obtain the appropriate ethical approval in order to use identifiable personal data.

8. What type of information will you be processing (please select all that apply)?

- No confidential or personal data
- Personal information, including signed consent forms

Event registration and consent forms will be stored in secure UoM sharepoint sites or in secured storage on campus, with identifying information deleted as soon as possible after it is no longer needed (e.g. participant emails will be deleted after the last event email is sent out).

Any other data will be anonymised and stored in the GitHub repo as well as UoM sharepoint sites.

9. How do you plan to store, protect and ensure confidentiality of any highly restricted data or personal data (please select all that apply)?

- Store data on University of Manchester approved and securely backed up servers or computers
- Store data in buildings, rooms or filing cabinets with controlled access

10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?

- No

11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?

- No

12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- Not applicable

13. Are you planning to use the personal information for future purposes such as research?

- No

14. Will this project use innovative technologies to collect or process data?

- No

15. Who will act as the data custodian for this study, and so be responsible for the information involved?

Dr. J. Kasmire will work with Gillian Meadows (UKDS PS admin) to ensure all data is stored appropriately.

16. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

2024-01-19

Assessment of existing data

Provide an explanation of the existing data sources that will be used by the research project, with references

Not applicable. No existing sources.

Provide an analysis of the gaps identified between the currently available and required data for the research

Not sure how to answer. This is a first move into this space so there is no data currently available.

Information on new data

Provide information on the data that will be produced or accessed by the research project

Anonymous data will be collected on workshop participants experiences with reproducibility, any barriers they are aware of that make reproducibility more difficult for some researchers, and any possible solutions to those barriers. The data will be collected through mentimeter polls/short answers/Q&A/etc. and through notes taken by the workshop leaders during the workshop.

Quality assurance of data

Describe the procedures for quality assurance that will be carried out on the data collected at the time of data collection, data entry, digitisation and data checking.

The workshop leaders will each take notes during the workshop and we will work together to consolidate and cross-check the notes after the event.

Backup and security of data

Describe the data security and backup procedures you will adopt to ensure the data and metadata are securely stored during the lifetime of the project.

The registration data will be stored securely in UoM restricted sharepoint sites until after the event. The hard copy of registration/consent forms will be kept on campus in a secure location for

approximately a year for auditing purposes.

The anonymous data collected during the event will be stored on UoM Sharepoint sites and in the UKDS GitHub repository to allow appropriate data sharing.

Management and curation of data

Outline your plans for preparing, organising and documenting data.

This is not really applicable to this data collection process. We are merely exploring ideas that those who are interested in reproducibility might have about reproducibility, barrier to reproducibility and potential solutions to those barriers.

We will share the anonymous mentimeter reactions and any notes taken during the workshop on our GitHub repository, but the real value will be through published research (also shared through the GitHub repo).

Difficulties in data sharing and measures to overcome these

Identify any potential obstacles to sharing your data, explain which and the possible measures you can apply to overcome these.

No anticipated difficulties.

Consent, anonymisation and strategies to enable further re-use of data

Make explicit mention of the planned procedures to handle consent for data sharing for data obtained from human participants, and/or how to anonymise data, to make sure that data can be made available and accessible for future scientific research.

The data collection purpose of the event is clearly spelled out in advertising. Attendees will be asked to confirm their consent that when they check in at registration. All data to be shared will be totally anonymous and non-personal.

Copyright and intellectual property ownership

State who will own the copyright and IPR of any new data that you will generate.

I guess the UoM and UKDS will own the copyright.

Responsibilities

Outline responsibilities for data management within research teams at all partner institutions

The lead investigator (J. Kasmire) will be responsible for ensuring that all anonymous data collected through the event is loaded to the appropriate GitHub repository. Gillian Meadows (UKDS PS admin) will ensure that personal registration data is stored securely on UoM sharepoint sites until it is no longer needed and then destroyed.

Preparation of data for sharing and archiving

Are the plans for preparing and documenting data for sharing and archiving with the UK Data Service appropriate?

No. Not applicable.

Is there evidence that data will be well documented during research to provide highquality contextual information and/or structured metadata for secondary users?

Yes. This is the whole point of the event (reproducibility) so detailed descriptions will be provided on the GitHub site and in any published articles.