FSS Data Organization Guidelines: Simple Quantitative

0. General information

0.1 Document version & date

Version 0.1
Date: 19/6/2023

0.2 Project title

Simple quantitative research project using Research Drive and Yoda

0.3 Project summary

This use case describes a fictional project using survey data. It uses Yoda for archiving only; for storage of data the project uses Research Drive. While Yoda is suitable for storage as well, Research Drive offers more fine-grained access control, which this fictional project needs to make sure student-assistants can’t access all information stored on the research drive.

0.4 At which VU Faculty is this project situated?

- Faculty of Social Sciences (FSW)

0.5 Your contact details

k.leuveld@vu.nl

0.6 List other people involved, including those at partner organisations in the project (if applicable)

Question not answered.

0.7 Funding organisation & grant number (if applicable)

Question not answered.

0.8 Project code (if applicable)

Question not answered.

0.9 Consulted data management expert(s)

Question not answered.

1. Data description

1.1 Will you collect and/or process personal data in this project?

- Yes

1.2 Will you use existing data? If yes, what is their source?

No

1.3 Will you collect or produce new data? If yes, please describe how.

Yes. Data is collected through a Survey.

1.4 Describe the population/participants/subjects that will be studied

Question not answered.

1.5 Do you process any of the following (personal) data?

Question not answered.

1.6 Do you process the personal data based on informed consent?

- Yes, using digital consent

1.7 On what legal ground will the data processing take place if it is not based on informed consent?

- Not applicable, I use informed consent

1.8 Does the data collection include any of the following types of personal data?

Question not answered.

1.9 If your research involves special categories of personal data (previous question) and you will not use explicit informed consent, what is the legal ground for the exemption?

Question not answered.

1.10 What kinds of outputs will you produce in this project? Please describe these data assets.

	Data Asset	Descripion	Format
Raw Data	Online survey data	Survey data hosted on qualtrics	N/A
	Exported survey data	Survey data exported to CSV	CSV
Processed Data	Pseudonymized survey data	Survey data without any identifying information such as names, IP addresses etc.	CSV
	Cleaned survey data	Pseudonymized data that has been cleaned and processed	RDS
Analyzed data	Tables and figures	Outputs obtained from statistical analysis based on the processed data	Word, PNG
Other	Analysis scripts	Scripts used to process and analyze the data	R
	Documentation	Questionnaires, proposals, Data Management Plans, ethics documents	Word

1.11 How much digital data storage will your project require?

- 0 - 50 GB

1.12 Will you collect physical data? If yes, please describe these.

No

1.13 Will you take measures to ensure data quality? Please describe these, if applicable.

Inputs to the survey will be validated upon entry. Answers will be further checked for inconsistencies after pseudonymization, and inconsistent answers will be cleaned.

2. Legal and ethical requirements, codes of conduct

2.1 What legislation applies to your research project? Please tick the relevant boxes for your project.

- General Data Protection Regulation (GDPR)/ Algemene Verordening Gegevensbescherming (AVG)

2.3 Do you require approval of an ethical committee for this project? If yes, please indicate which ethical committee and whether you have obtained approval for this project.

- No

Self-check number 2023-06-8686

2.4 Will you work with data for which intellectual property and/ or confidentiality are an issue? If yes, please describe.

- No

2.5 Do you plan on generating a marketable product from your research project? if yes, please describe

- No

3. Storage and back-up during the research process

3.1 What measures will you take to secure and protect data during the research process? Please describe, for each separate data asset you described for question 1.10, how you will ensure data security, where the data assets are stored & backed up, and who has authorization to access the asset.

	Data Asset	**Storage & Backup**
Raw Data	Online survey data	Qualtrics
	Exported survey data	Yoda
Processed Data	Pseudonymized survey data	Research Drive
	Cleaned survey data	Research Drive
Analyzed data	Tables and figures	Research Drive
Other	Analysis scripts	Research Drive
	Documentation	Research Drive

Security measures:
- Qualtrics: Qualtrics is only accessible to authorized users using institution sign-in. Data will be removed once the survey is complete.
- Yoda: raw data on yoda will be archived as soon as pseudonymization is complete so that data can no longer be changed. The archived data is only accessible to authorized personnel using multi-factor authentication.
- Research Drive: research drive is accessible to authorized users. Users are only given access to the folders they need for their activities in the project.

3.3 Which tools are used in the collection, processing or storage of data during research?

- Research Drive (Surf)
- Qualtrics
- R (software) *
- Yoda

3.4 What other tools or software do you intend to use during your research?

Question not answered.

3.5 Is it necessary to transfer the (physical or digital) data assets to other locations or research partners? If yes, please describe how you secure the file transfer.

- No

3.7 Do you transfer personal data outside of the European Economic Area (EEA)? If Yes, please provide additional information

- No

4. Data archiving and publishing

4.1 Which data assets will be archived and which will be published?

	Data Asset	Archive or publish
Raw Data	Online survey data	Deleted
	Exported survey data	Archived
Processed Data	Pseudonymized survey data	Deleted (can be derived from raw data using scripts)
	Cleaned survey data	Deleted (can be derived from raw data using scripts)
Analyzed data	Tables and figures	Published
Other	Analysis scripts	Published
	Documentation	Published

4.2 Where will you archive your data assets?

- Yoda

4.3 What other archive(s) do you intend to use to archive data assets?

Question not answered.

4.4 For how long will the data be available in the archive?

10 years after publication.

4.6 Where will you publish your data assets?

Yoda.

4.8 How will you ensure your data assets get a persistent identifier (e.g. a DOI-code)?

Yoda mints a DOI upon publication of the metadata.

4.9 Will you register your datasets in an online registry other than PURE? If yes, where?

No.

4.10 Are there restrictions to data publishing? If yes, please specify the reasons and list the data assets you do not wish to share publicly.

The data is sensitive, and to protect the privacy of the data subjects, it will not be published.

4.12 When will you share the data? If not immediately after completion of the project, please specify the reasons.

Immediately.

4.13 Please indicate the license and/ or terms of use under which you share your data.

All published materials will be available under a CC-BY license.
The archived data sets will only be shared for verification purposes under a strict data sharing agreement limiting the use of the data to the verification.

5. Documentation

5.1 What documentation and metadata will accompany the project?

The project is documented by the project proposal, data management plan, and documents relating to ethcis clearance.

5.2 What metadata and documentation will accompany the data assets?

The raw data will be accompanied by a codebook describing the variables, as well as a text-version of the questionnaire.

5.3 What methods, software or hardware are needed to access and use your data?

The data is stored in CSV, which is an open format that can be universally read.

6. Data management responsibilities and resources

6.1 Who will be responsible for management of the data assets during the project? Please specify their name, position, role in the project, and faculty/ institution/ group.

6.2 Who will be responsible for management of the data assets after completion of the project (e.g. the project lead/ dedicated data manager/ department head)? Please specify their name, position, role in the project, and faculty/ institution/ group.

Department Head

6.3 For data that are only available upon request, what methods will be used to handle requests for access and how will data be made available to those requesting access?

The data will be archived on Yoda, which handles this.

6.4 What resources (for example financial and time) will be dedicated to research data management? Please estimate their cost.

The project proposal dedicated two hours per week to data management, as well as 1000 euros in unforeseen costs for additional storage or software.

The tools used and advice of the data steward are provided by the faculty without cost.