Plan Overview

A Data Management Plan created using DMPonline

Title: (Scoping Work: ESRC IAA) - Community Partnerships For Water Resilience

Creator: Julian Dyer

Principal Investigator: Dennis Onyango Ochieng, Ellen Dyer, Julian Dyer

Affiliation: University of Exeter

Funder: UKRI Future Leaders Fellowships

Template: University of Exeter

ORCID iD: 0000-0003-2954-9294

Project abstract:

This project will explore how to create resilient boreholes and other infrastructure to create water security in Turkana, Kenya.

Boreholes and water projects in Turkana, Kenya suffer from alarming failure rate - 60% within two years according to UNICEF audit. This is likely due to institutional failures and the political economy of projects build and designed by NGOs, then handed over to local governments and communities to maintain. Without ownership, capacity, or buy-in, the maintenance doesn't get done. This field experiment in partnership with the major funders of boreholes in Turkana will test various models of community partnerships to fix this problem.

The scoping work for this project aims to build partnerships and identify a feasible approach for a joint project. We will interview stakeholders from projects elsewhere in Kenya that have tried different models of community partnerships, and will use these interviews to inform the stakeholder engagement. We will also interview successful and unsuccessful water projects in Turkana to learn more about the reasons for failure to facilitate our partnership building.

ID: 155223

Start date: 01-08-2024

End date: 01-10-2024

Last modified: 20-01-2025

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

(Scoping Work: ESRC IAA) - Community Partnerships For Water Resilience

Data

If you are re-using existing data, what licences or terms of use will you have to comply with?

No existing data.

How will new data build on and relate to existing data?

No existing data.

What types of new data will you create and in what format?

At the scoping phase, we will create interviews and qualitative data collected from semi-structured questionnaires stored as text documents initially then collated into a SQL database.
In the intervention design phase, we will use surveys and lab-in-the-field methods to measure the preferences of participants over different types of intervention. This will be collected using tablets and stored as encrypted quantitative data on secure servers located in the EU.
We will collect similar survey data on behaviour, economic outcomes, beliefs about climate shocks, and attitudes towards water projects and payments using tablets stored on the same secure servers.
We will take photos of the water projects being discussed and their usage in households and other public institutions. Only adults will be included in photos used in the study.

Can you estimate the size of the data you will create?

At the scoping phase this will be quite small as we are considering less than 50 interviews, certainly under 50mb.

What methods will you use to capture your data and how will these ensure that your data are high quality?

For scoping, semi-structured interviews conducted by an experienced researcher, partly with key informants, partly with focus groups, as well as household interviews with end-users of water projects.

Documentation and description

What contextual information is needed for you or someone else to understand your data?

As this data will be in the form of semi-structured interviews, the only contextual information required will be basic understanding of the climate risks faced by pastoralist communities in Northern Kenya, as well as basic information on the design and use of water boreholes in Turkana.

Photographs will require no additional contextual information to understand.

How will you capture contextual information?

For qualitative scoping work, where any important contextual information is required, the person conducting the interview will include this as endnotes to the interview transcript.

Will you use any metadata standards?

Not necessary for scoping data.
For the main experiment, we will use a readme file to ensure all the different piece of data and documentation are properly stored and understandable.

Data Protection

Where will you store your data and how will you ensure that they are backed up? Will you use University-managed data storage or will you need to set up your own back-up procedures?

All data will be stored on a secure Sharepoint site that has been created for this project on the University of Exeter Sharepoint site. We will also archive this data separately on the Uni Exeter Research Data Storage system.
Only relevant members of the research team will be able to access this data to ensure personally identifiable information is protected. Analysis will be performed on cleaned, anonymized, data. All photographs will be edited to blur faces so no individuals will be identifiable.
Survey data will be encrypted when stored on the ONA Data servers.

How will you secure your data? What methods will you use to restrict access to your sensitive data? Will you encrypt hardware when working off campus?

At the scoping stage, sensitive data will not be collected, and all data we collect will be stored on a secure Sharepoint site, with any physical copies of transcripts destroyed after interviews are transcribed and uploaded. Access to raw interview transcripts with identifiable information and original photos will be limited to research team members. All photographs will be edited to blur faces so no individuals will be identifiable.
For the main experiment, all personally identifiable data will be stored on password protected Sharepoint sites, and should data need to be stored temporarily on portable devices they will be encrypted. We will securely share data among collaborators by managing access to the project Sharepoint page and managing access to the data on the ONA Data server.

How will you protect your research participants? Will you obtain informed consent for data retention and sharing? How will you anonymise data to safeguard the privacy of your participants?

We will obtained informed consent including details on data retention and sharing. All information shared as part of this data will have names and any other identifiable information removed before being shared. We will anonymise data by removing all personal identifiers and obscuring geographic information to an appropriate level where individual households cannot be identified. All photographs will be edited to blur faces so no individuals will be identifiable.

Retention and preservation

Which subsets of your data will you keep at the end of your project? Will you retain anonymised versions but destroy personal data and identification keys? Will you retain all of the raw data or is a processed version more suitable to preserve? Do you need to keep all intermediary files or would you only need to refer back to input files or a final version?

At the end of the scoping phase of this project we will only keep the anonymous sections of interview transcripts we include in research publications or other outputs. We will only keep fully-anonymous data with all identifiable information removed. No intermediate or original files would need to be kept. We will only retain the anonymised photographs.
For the main experiment, we will keep the survey data, retaining anonymised information without identification keys beyond the end of the project. We will keep anonymised intermediary files for completeness.

How will you prepare your data for long-term preservation? Are you able to convert your data to open file formats? What contextual information do you need to retain so that your data remain understandable and usable?

Interview transcripts will be stored in open file formats and will not require any contextual information beyond the notes included by the enumerator.
We will delete all personally identifiable data from the dataset and store it in standard CSV and STATA-dta formats on the PI's website and on the journal website when it is published. We will retain the survey manuals and questionnaires to ensure it is understandable and usable. We will delete the original identifiable photos.

Where will you archive your data to ensure that they are preserved and sustained for several years after your project ends? Will you submit your data to a specialist data repository/centre and if so, have you consulted them about your requirements?

We will use the University of Exeter's institutional repository and the Research Data Services system for secure long-term storage.

How big will your final dataset be and will there be any costs associated with archiving them, such as data deposit charges?

The dataset will be quite small as it is only a small number of interview transcripts and will not require any costs to archive, and a larger number of quantitative surveys that are not particularly large. We do not anticipate any costs to archiving this data.

Data sharing

Can you demonstrate that you'll plan ahead to maximise data sharing? For example, will you only share a subset of the data where informed consent was granted for data sharing?

We will only share identifiable information within the research team, and only fully-anonymised portions of the interviews will be included in any external communication where participants have consented to this.
We will include data sharing as part of the consent process for all observations to demonstrate our commitment to open data.

Are there any reasons why you would not be able to share some of your data? Would they be covered by data protection legislation, licence restrictions, or contractual confidentiality clauses? Are there ethical reasons why your data should not be released?

There are no reasons not to release the fully-anonymous interview data from scoping phase.
We will not share the exact geolocations of households from surveys, which will be helpful for analysis as it will allow us to compute distances to various key items of interest. We will not share exact geolocations for privacy and ethics concerns, and will instead share obscured geographic data that cannot be used to identify households.

When will you share your data? Will data be made available upon first publication of findings or within a limited period after the end of the project? Do you need to delay publication to allow for commercialisation or patent applications? Will you embargo your data to allow for a limited period of exclusive use?

We will share the relevant segments of the interviews as quoted within the research publications. We will not need to delay sharing of this.
We will share data upon first publication of findings, with no reason to delay beyond this point. We will not embargo the data beyond the date when the relevant data is used in publication.

How will you disseminate your research? Will you include a data access statement in published articles? Does your chosen method of data preservation provide a persistent identifier such as a Digital Object Identifier? What licences will you assign to your data?

We will disseminate research via academic publication. Data access statements and DOIs will not apply to the segments of interviews included in publications and grant applications.
The institutional repositories we will use at the University of Exeter will have a DOI that can be shared.

Data Protection Impact Assessment

Will your research involve human participants or personal data?

Examples of personal data could include names, addresses, photos, video, ID numbers, DNA, IP addresses, job titles etc.

- Yes

What do you require this personal data for? What is the purpose of using the personal data?

For scoping work, we will not keep this personal data as part of the data beyond identifying key demographic characteristics. Personal data will not be used in the analysis or stored or shared.

For the main experiment, we will need personal data to track participants in order to implement treatment and find them for follow-up surveys.

How are you making people aware of how their personal data is being used? Do you need to update your privacy notice?

Participants will be informed that their personal data will not be shared and all information shared beyond the research team will be anonymous.

Which conditions for processing apply for your project? For Special Categories please ensure you select at least one from Section 1 and one from Section 2 below. Please select all that apply and provide any additional details.

Section 1: Conditions for Personal Data

- **The data subject has given consent to the processing (please provide the consent wording and where it is stored)**
- **Contractual necessity (please confirm which contract this relates to)**
- **Compliance with any legal obligation (please document which legal obligation)**
- **To protect the vital interests of the data subject (please provide details)**
- **Functions of a public nature or task in the public interest (please provide details)**
- **Legitimate interest of the Data Controller (please provide details of legitimate interest)**

Section 2: Conditions for Special Categories Data

- **The data subject has given explicit consent to the processing**
- **Necessary so that you can comply with employment law**
- **To protect the vital interests of the data subject or other person**
- **The processing is carried out as part of the legitimate activities of a not-for-profit organisation**
- **The individual has deliberately made the information public**
- **The processing is necessary in relation to legal rights**
- **The processing is necessary for administering justice or for exercising statutory or governmental functions**
- **The processing is necessary for medical purposes**
- **The processing is necessary for monitoring equality of opportunity**

Participants will fill out a consent form giving consent to the collection of interview orsurvey data and the sharing of anonymised information.

Is all the personal data you are using necessary? Are you collecting enough to carry out the work, is there any you could do without to limit the risks to the individuals?

Yes.

How are you ensuring that personal data obtained from individuals or other organisations is accurate? How will you keep it updated?

We will be collecting information directly from participants, and there is no need to ensure accuracy beyond this. We will not need to keep it updated for this stage of the project.

How long will you keep the data and how will you dispose of it? Are the retention periods on the University Retention Schedule?

We will dispose of identifiable data after project completion. We will store anonymised data long-term in keeping with University of Exeter policy.

Where will the data be stored? If storage is in the cloud, where is the physical server? Will you need to transfer the data outside the EEA? If yes, how will you ensure adequate protection?

All data will be stored securely on Sharepoint during analysis, with a backup on the Research Data Services system, before being archived on the Uni. Exeter's ORE repository.

Will you be able to meet all the Data Subject Rights? Can you provide copies of data if requested? Are you able to fully delete the data (not just archive)?

Yes.

Please briefly document below any risks with the use of personal data and how you will control such risks. Include technical controls (IT security, encryption etc), physical controls (location, locked room etc), personnel controls (training, access control etc), and procedural controls (contract, polices etc).

We will control these risks by storing all data on a secure Sharepoint site that we regularly monitor to maintain secure access by team members only. We will encrypt all survey data when it is being transferred to the ONA Data survey server platform and delete data from the tablets on which it is collected after upload. All photographs will be edited to blur faces so no individuals will be identifiable.

We will only store data temporarily in encrypted form on portable devices, which will be kept in locked drawers in locked officers when not in use.