**Tips for Preparing an Application for the Ethics Committee:**

**Guidelines for Data Management**

**Zsofia Garai-Takacs (Ethics Lead in Clinical Psychology)**

**Reviewed by Simon Smith (Research Data Support Officer) and the Ethics Committee of the School of Health in Social Science**

As a student/staff member you should have already completed the mandatory data storage trainings such as MANTRA (especially relevant here is the module on ‘Keeping research data safe’) and the Data Protection and Data Protection for Researchers on Learn. The present guide is to provide tips and examples of good practices from the Ethics Committee.

1. **Identifying and non-identifying data**

When thinking about data management and storage it is important to make a distinction between data that is not identifying (e.g. someone’s scores on a psychological scale) and identifying data such as name, address, phone number, email address, photos of people and voice or video recording.

If you need to collect any identifying data, make sure to store it separately from any research data, that is, in different folders or even on different servers and avoid connecting the identifying data to the research data. For instance, collecting participants’ email addresses to include them in a prize draw or because they requested a summary of the results of the research project does not necessitate connecting the email addresses to the research data you collected about them (e.g. their responses to a survey, their scores on a psychological test or their interview transcript). It is also a requirement to store identifying data in encrypted folders (for guidance see <https://infosec.ed.ac.uk/how-to-protect/encrypting>).

In some cases researchers do need to connect participants’ identifying information to their research data because of repeated measures or needing to match data of dyads (e.g. parent-child or spouses), for instance. In these cases store the sheet connecting the identifiable information to the participant IDs in a separate file, away from the research data. Also, it might not be necessary for the whole research team to have access to the spreadsheets including identifying information, give some consideration to who should and should not have access. It should be noted, however, that even in the case of repeated measures and when you need to match the data of dyads, there are techniques to avoid using identifying information for these purposes. For more information:

<https://journals.sagepub.com/doi/full/10.1177/0013164419843576>

<https://www.sciencedirect.com/science/article/abs/pii/S1054139X17300253>

<https://www.sciencedirect.com/science/article/abs/pii/S0149718919304914>

If you need to collect any identifying data, make sure to store it only as long as absolutely necessary. For instance, if you are collecting participants’ email addresses in order to enter them in a prize draw, please make sure to delete the email addresses right after said prize draw. Similarly, if you need to have a spreadsheet connecting participants’ IDs to identifying information, make sure to delete this spreadsheet as soon as it is no longer necessary (e.g. upon completion of the last wave of data collection in a longitudinal study). Also, when collecting audio or video recordings of interviews or focus groups, these should be transcribed as soon as possible with the recordings deleted right after. It is good practice to do this within 1-2 weeks. If you need to store these longer, please provide the reasons on your ethics application. In either case, make sure to explicitly state in the participant information sheet and consent form how long you will store identifying and anonymised data.

It is a GDPR requirement to delete personal data securely. Instead of simply hitting the delete button, it is recommended to encrypt the data before deleting and then delete it. It is a good idea to set the encryption key to something you will not remember (e.g. random bash on the keyboard).

In regards to recording audio or video online, the university recommends using Teams, which is GDPR compliant, or applying OBS recording (available from the UoE Software Centre) if these happen on other platforms such as Zoom. This is because if you use OBS recording, you can save the recordings straight to DataStore. In regards to transcription, using the transcription function of Teams is fine. Researchers are encouraged to move recordings to secure storage as soon as possible. Alternatively, researchers can also use the transcription function in MS Word. In case you would like to use third-party services to transcribe recordings, make sure that the company is fully GDPR compliant and that there is a contract in place.

After participating (e.g. finishing an interview, a survey or a lab visit) participants can only withdraw their data as long as it is not anonymised. Alternatively, you can share a code that is connected to their participant ID with the participant in order for them to request their data to be deleted after participating. That is, unless you share such a code with participants, you will not be able to delete their data after the interviews were transcribed and anonymized, after a response was submitted to an anonymous survey and after you deleted all identifying data following a lab visit. Participants should be informed of this in the participant information sheet.

1. **Where to store data?**

**Paper-based data:**

 Make sure to digitize the data as soon as possible and follow guidelines regarding where to store data digitally below. This could include, for instance, scanning paper questionnaires or paper notes. In the meanwhile, paper-based data should only be stored in a locked cabinet on university premises.

If paper-based consent forms are used, researchers are required to store those as long as the data is stored.

**Digital data:**

 Data should be stored on one of the university’s secure servers. For non-sensitive, anonymized data, OneDrive is okay, however, consider using DataStore, especially for sensitive and/or identifying data. While collecting the data, Qualtrics, JISC and Teams are fine, however, it should be transferred to a secure server of the university once data collection is finished. When working off-campus, please make sure that your laptop is encrypted and you use VPN.

 For more details on the university’s secure servers: <https://www.ed.ac.uk/sites/default/files/atoms/files/quickguide_03_proof02_0.pdf>

The Research Data Support Team provides support for everyone to have a Data Management Plan (<https://dmponline.dcc.ac.uk/>).



**3. Where *not* to store data?**

When storing data please make sure to use the university’s servers and not Google Drive, Dropbox or any similar services. Data should not be stored locally on computers, laptops or external drives (e.g. pen drive) either. Also, you should not use AI-driven transcription and translation services. They are absolutely not GDPR compliant, regardless of what is stated on their websites.

1. **Sharing and receiving data**

To share any identifying information, you need participants’ consent to do so. When only sharing anonymised data it is still good practice to inform participants in the participant information sheet. For instance, you can include a statement that the anonymised data may be used for future ethically approved studies or that the anonymised data may be shared with other researchers via online data repositories.

-If working with external collaborators, DataSync can be useful (and used even for identifying data if needed). Consider a data use agreement with external collaborators.

-It is good open science practice to share anonymous data publicly (e.g. on Edinburgh DataShare, UK Data Service or OSF) upon publication of a study. The university’s open data archive is the Edinburgh DataShare.

-If you cannot share the data openly, you can create a dataset record in Pure and archive data on DataVault.

-If you would like to receive research data, even if anonymous, you need to submit a Level 1 application to the Ethics Committee.

1. **How long should I retain the collected data?**

Data retention should be included in the data management plan (DMP) of your research proposal and based on requirements by your funders and data providers, as well as disciplinary guidance. A suggested minimum period of retention of three years from the end of the project is recommended by the Research Data Management Policy at the University of Edinburgh.

For further information/advice:

Research Data Service:

<https://library.ed.ac.uk/research-support/research-data-service>

Contact: data-support@ed.ac.uk

MANTRA:

<https://mantra.ed.ac.uk/>

Quick guides:

<https://www.ed.ac.uk/information-services/research-support/research-data-service/guidance>

Sensitive data:

<https://www.ed.ac.uk/information-services/research-support/research-data-service/during/sensitive-data>

Guidance for students:

<https://data-protection.ed.ac.uk/guidance/personal-data-processed-students>