Guidance on Research Data Handling for UG and PGT Students[*]

This document is intended to assist in completing the research data management section of the SSP Application for Research Ethics Approval for **UG and PGT students only**.

If you/your supervisor believe that the project work may lead to an academic publication, then please use the 'PGR and Staff' version of these guidance notes.

Note that this document focusses on the typical research projects in SSP. Some other types of research (e.g. involving NHS staff, patients or patient data, or animals) may have their own specific guidelines.

Will **personal data** (including photos, video or audio) be collected, recorded or used?

In general, personal data is data from or about a living individual, who might be identifiable from that data (perhaps when combined with other available data). There is a further guidance on identifying personal data on the SSP-REC website.

> This may include data that you capture, such as video or audio recordings.
>
> It may also include data that you get from elsewhere, from which an individual might be identified. For example, you still need to consider the ethics of using data obtained from social media sites or other researchers.

What data (or materials) will be collected or used?

Identify the type of data and also the subject or content of the data.

> E.g. questionnaire responses, logs of network traffic from a home network, field notes of direct observations of…, audio recordings of interviews. Note: video recordings of interviews are not permitted.

What if any **constraints** apply to use of this data (or materials)?

In general participants should give (or have already given) informed consent to use of their data that includes what you propose to do.

> If the data was collected by someone else, then check what consent was given when it was collected.
>
> If the data is obtained from a third party such as a social media site, then it will typically also be subject to a license agreement that restricts what can be done with it.

Some data may also be subject to copyright, e.g. the words that people say or write. You need to ensure that your use also respects this and/or permission has been obtained.

[*] Adapted from guidance by Chris Greenhalgh, School of Computer Science - Updated Sept 2020

Some data may also be company confidential, e.g. internal details of a business process or commercial operation, obtained under a contract (e.g. a non-disclosure agreement). Again, note these constraints if they apply (typically your supervisor will be able to advise in this case).

How will this data (or materials) be **collected or obtained**?

Explain briefly how the data will be obtained. For example, will you collect or record it yourself? Where/when?

Consider whether this raises any additional complications, e.g. might you also record other people who are not part of your research and who haven't given consent?

Note that for online surveys you still need to comply with GDPR/data protection requirements. For simple surveys you should be able to use the University-provided Office365 'forms' functionality.

> Try to make survey responses anonymous if possible and include information about the study at the start of the survey with a compulsory checkbox to say that they understand and agree to participate.

How will this data (or materials) be **processed before analysis**?

What if any processing will be done on the data before it is analysed?

> For example, audio recordings may be transcribed.

Also make clear whether/what anonymization takes place at this stage.

> For example, use pseudonyms or IDs for participants, not their real names; removing identifying information from free text responses.

How will this data (or materials) be **stored and secured** (a) **during the research** (b) **after the research**?

Please use the University-provided Office 365 OneDrive to store all research data, including personal data (the capacity is up to 5TB/user). You can share this with your supervisor as appropriate.

> Note that storage in OneDrive is linked to the individual who created it, and when you leave the University it will be deleted.

If your data includes personal data or commercially sensitive data then this is classified as 'Restricted' Data, and the following apply:

> Restricted data must be secured appropriately, e.g.
>
> > Laptops and mobile devices must be protected at all times, e.g. encrypted and information secured with One Drive.

Any portable storage must have a secure physical backup, e.g. on OneDrive.

Also note that OneDrive cloud storage is encrypted, but if you download files (including using OneDrive Sync) to a local machine then you need ensure that the local storage is also encrypted. For example, enable disk-level encryption on your computer, transfer files in encrypted archives and/or encrypt individual files (e.g. using Office's built-in encryption support).

Any sensitive personal data (e.g. health, sexuality) MUST be encrypted (just being stored on a 'secure' store is not enough).

Where personal data cannot be encrypted (e.g. as it is being captured on a video camera or voice recorder) you should plan to transfer it to a securer form of storage as soon as possible.

How will the data be **formatted**, i.e. what data formats or encodings (e.g. file types, media encodings) will be used?

This is not a major issue for student projects, but in general consider using standard and open formats.

How will the data be **organised**?

E.g. what folder structure will you use, what metadata will you provide?

It may be useful to subdivide the data into:

1. **Method and approvals** (University or Public) – e.g. ethics forms, information sheet, etc. These will NOT usually be restricted data.
2. **Participants** (Restricted) – i.e. completed (scanned) consent forms, participant lists, records of withdrawals (and steps taken). These will be restricted data.
3. **Raw Data** (Restricted) – i.e. sensitive or personal data such as audio recordings, logs with location or physiological data, as captured. There may be several sub-folders, e.g. for data type or session. Details of anonymization should be recorded here (if relevant).
4. **Working Data** (Restricted, University or Public) – i.e. data as used for the main research activity, anonymised where possible. There may be several sub-folders, e.g. for data type or session.
5. **Analysis and Results** (Restricted, University or Public) – i.e. work being done on/with the data. There may be several sub-folders, e.g. for particular analyses.

(Omit sections that are not relevant. Research projects may also include: 6. Archive Data and 7. Published Data.)

Where there are different parts to the research (e.g. related studies) it is probably better to reflect this at the top level and use the above sub-divisions within each sub-study. Where data from multiple studies is combined, identify this as a separate top-level activity.

Typically, each data folder should basic metadata, i.e. a short description of the data that it contains, source, subject and format(s). For example, this might be in a textual 'README' file in each directory.

How will this data (or materials) be **analysed**?

What kinds of methods will you use to analyse the data? What tools will you use?

> You don't need to go into lots of detail but try to make clear whether the results will potentially include personal data or not. For example, statistics from a population will generally not be personal data, but fragments of video featuring recognisable individuals will be.

How will this data (or materials) be **reported** in publications?

This refers to what will appear in your dissertation or other reports. You don't need to give a lot of detail but make it clear what kinds of results might be published and how personal these might be, as above.

In most cases this will be anonymous.

> Where users are referred to specifically, pseudonyms or study-specific identifiers will normally be used.
>
> If users are potentially identifiable then they will have explicitly agreed to this (and in some case, e.g. artists or public figures) may have explicitly requested it.

Note that:

> For taught modules the coursework submission will usually only be seen by the module convenor, second marker(s) and external examiners. In some cases, students may see each other's submissions; check with the module convenor in this case.
>
> For undergraduate projects normally the supervisor, other markers and supervisors and the external examiners will see the submitted dissertation. A few excellent dissertations are made available to other students (if the student agrees to this).
>
> Masters projects might be submitted to the University's dissertation archive, in which case they will be publicly accessible.
>
> If you/your supervisor believe that the project work may lead to an academic publication, then please use the 'PGR and Staff' version of these guidance notes.

How and **when** (if ever) will this data (or materials) be **reused**

The focus here is on re-use by you or your supervisor. Be realistic: most data from student projects will never be reused.

How and **when** (if ever) will this data (or materials) be: **archived, indexed, published** or made available to others?

The focus here is on long-term and discoverable preservation of research data beyond the initial research project. Again, most data from student projects will never be of interest to another researcher, so there is no point archiving or publishing it. If there is then please refer to the version of these guidelines for staff or PGR research projects.

How and **when** (if ever) will this data (or materials) be **deleted or destroyed**?

Note: the information in this section is provisional and subject to amendment.

You should retain (or be able to reproduce) the data underlying a coursework or dissertation at least until the work has been assessed and ratified by an external examiner, in case any concerns are raised about academic misconduct, i.e. until you have received you confirmed marks for that module after the corresponding external examiner's meeting (typically in June [UG] or October/March [PGT]).

> If you use your University OneDrive storage, then the data will normally be deleted after you finish the qualification for which you are currently registered.

Note: If you/your supervisor believe that the project work may lead to an academic publication then the requirements for retention of data are different – see the 'PGR and Staff' version of these guidance notes.

If human subjects are involved, then at what point(s) can they **withdraw** and what will happen in each case?

In general, you won't be collecting any data from the moment the person withdraws.

In general, it is best practice to allow retrospective withdrawal, i.e. to remove as far as possible data captured or provided prior to their withdrawal.

What will happen to this data if/when you **leave the University**?

In general, it will be deleted. But if you/your supervisor believe that the project work may lead to an academic publication then the requirements are different - please use the 'PGR and Staff' version of these guidance notes.