University of Reading

DMP Guidance

For participant-based research

Introduction

This guidance accompanies the DMP Template for Participant-based Research. The template can be used to plan for the management and sharing of participant-based research data in compliance with University research ethics and data protection requirements.

You should complete the DMP with reference to the guidance below and the Data Protection for Researchers guide provided by the Information Management and Policy Services (IMPS) office.

Applications for ethical approval submitted to the University Research Ethics Committee must be accompanied by a DMP prepared using this template. All sections of the DMP must be completed; is a section is not applicable, enter N/A. The DMP will be reviewed by the IMPS office and the Research Data Manager, and you will receive feedback on the DMP when the UREC opinion is given.

If you have questions about completing the DMP, please contact Robert Darby, Research Data Manager, r.m.darby@reading.ac.uk / 0118 378 6161.

Definitions

The DMP is concerned with how you will manage both personal data and confidential information in relation to research data.

Personal data is information as defined in data protection laws relating to natural persons who can be identified or are identifiable, either directly from the information, or indirectly from the information in combination with other information. Personal data must be processed in accordance with data protection laws. Personal data that contain sensitive information, such as details of a person's medical history, constitute a special category under data protection laws, and are subject to stricter processing requirements.

Confidential information is non-public information relating to an identifiable living individual or legal entity. Those with whom this information is shared have an ethical obligation to maintain the confidentiality of the information. In research, confidential individual information and personal data are often the same. Information may be confidential for other reasons, for example, if it is non-public information relating to a

business or other organisation. Examples include commercially confidential information, and sensitive location details (e.g. relating to endangered species, military sites, etc.).

Research data is information collected to answer a research question. Where the data have been collected from research participants, they are usually processed at some stage in the research to remove direct identifiers such as names and contact details, although they may still include indirect identifiers (such as key codes, age, job title, etc.). De-identified research data is generally suitable for public sharing, although care may need to be taken to unlink pseudonymous key codes and mask or remove indirect identifiers before making research data widely available.

Separating identifying information from research data during research is good practice, e.g. by keeping a separate locked database of participants, and using pseudonymous key codes to designate participants in research data files. But under data protection law pseudonymous data remain personal data if any key codes by which the data could be re-identified continue in existence, so they must still be treated with care.

Guidance on completing the DMP

1. What research data will be collected?

Describe:

- the types of data to be collected, e.g. interviews with farm workers about pesticide use; analyses of blood samples from trial patients; structural and functional MRI scans, with processed images, statistical data and analysis code;
- all media/formats data will be collected in, e.g. paper questionnaires; audio recordings; text transcripts; online tools with export to standard tabular formats; proprietary instrument formats;
- the anticipated quantity or scale of each type of data, e.g. 20 one-hour interviews; analyses of blood samples taken from 30 patients weekly over 8 weeks; MRI scan sessions of one hour's duration for 50 participants, generating in total ~ 1 TB raw and processed data.

2. What personal data and confidential information will be processed?

2.1 Personal data

Specify the personal data (as defined in data protection laws) that will be collected, e.g. name, address, email address, photographs in which individuals are identifiable, geolocation data, IP addresses.

A checklist of personal data is provided in the DMP.

2.2 Special category data

Specify any special category (sensitive) data you will collect, as defined in data protection laws, e.g. information about an individual's health, political opinions, etc.

A checklist of special category data is provided in the DMP.

2.3 Confidential information

Specify any confidential information not specified above that will be collected, e.g. non-public information relating to a business or other organisation.

Guidance

See the entries for Personal Data and Sensitive Personal Data in Data Protection For Researchers.

3. How will data be stored and transferred during the project?

3.1 Locations

Identify all locations where data and supporting materials will be stored, including:

- the primary storage location for project data and documents, e.g. a location on the University network or a University OneDrive account;
- instruments for data collection and analysis, such as online survey tools, analytic instruments, audio and video recording devices;
- working locations for data processing and analysis, such as project members' laptops and other devices;
- locations for storage of non-digital data, e.g. signed consent forms, paper questionnaires. Consider storage in the field/in transit as well as on University premises.

For each location, indicate whether it will be used to store/process identifying information or de-identified research data, and provide details of access controls that will be applied, such as password protection, or encryption of files or devices.

3.2 Risk management

Describe any administrative measures that you will take to control the risks of inappropriate disclosure of personal data/confidential information, e.g. storing participant records separately from research data in closed storage areas accessible only to authorised users, using key codes in research data to enable relinking of data to participants as required (pseudonymisation).

Describe procedures for secure transfer between locations as necessary, e.g. file encryption; download from cloud location via end-to-end encrypted channel; transfer to University OneDrive account, or to a University network location via VPN. Where materials are obtained in non-digital forms (e.g. signed consent forms), plans for safe storage and transport of these from collection locations to secure storage on the University premises should be addressed. If you intend to digitise and destroy hard copy originals, this should be stated.

3.3 Authorised persons

Specify who will be able to access the identifying information and how you will ensure they process the information securely, e.g. through training, supervision where

appropriate, and adherence to agreed protocols for accessing confidential information in secure environments. If you will be using a service supplier acting as a data processor, such as a professional transcription service, have the terms of service been discussed with and agreed by Procurement or your IT Business Partner?

Guidance

See the RDM web pages for guidance on data storage and information security, and online survey tools (bottom of page).

Guidance on pseudonymisation can be found in Data Protection for Researchers.

Guidance on data protection requirements for suppliers is provided by Procurement.

4. How will research data be preserved and shared on completion of the project?

4.1 Research data to be preserved and shared

Identify the research data that will be preserved and shared at the end of the project by deposit in a public data repository. If some or all research data will not be shared, explain why this is the case.

4.2 Preparation of data for sharing

Describe the measures that will be taken to ensure data are suitable for sharing, e.g. informing participants during recruitment that data will be shared; securing consent for data sharing; anonymising data prior to deposit/sharing; or depositing confidential data in a data repository under a controlled access policy.

4.3 Data repository

Identify the data repository or repositories that will be used to preserve and share data. If no data repositories will be used, explain what other solutions will be used to ensure research data are preserved and can be accessed publicly or on authorised request after the project.

Guidance

The University's Research Data Management Policy requires researchers to preserve primary data collected in support of project findings, and to make them accessible to others by deposit in a suitable public data repository, unless there is a valid legal, ethical or commercial reason for withholding access to them. The RDM web pages provide guidance on choosing a data repository.

Most research data collected from research participants can be safely and ethically shared once they have been anonymised (i.e. identifying information has been removed). It is not acceptable simply to state that research data cannot be shared for confidential reasons. If you do not intend to share data you must explain why they are not suitable for sharing.

Be aware that in order for publicly-disclosed data to be fully anonymised, any means of linking them to participant records stored internally should be destroyed. If you have used pseudonymous key codes in your dataset which are linked to internal participant records, the key codes in the public dataset should be replaced by random identifiers. Guidance on anonymisation can be found in Data Protection for Researchers.

Where confidential data cannot be rendered safe for public sharing (for example, where identifying information is intrinsic to the data and cannot be removed), it may still be possible to share them with authorised users on a restricted basis. Some data repositories provide controlled access procedures for managing safe access to confidential research data. Further information about such services can be found on the Where to archive data web page.

Consent is not required to share anonymised data, although as a matter of good practice research participants should always be informed of plans to make any data collected from them available to others.

You should in any case take care that your consent procedures do not preclude public sharing of anonymised research data. Do not set a time limit on the retention of the data collected from participants, or state that all data will be destroyed at the end of the project, or undertake that data will not be shared outside of the project. Such undertakings are not required by data protection law or research ethics policy, and they will prevent you from making your research data publicly accessible, even if they have been anonymised.

A sample consent form, with consent formulae for public sharing of anonymised data, and for restricted sharing of safeguarded data, can be found on the IMPS website. The UK Data Service provides guidance on seeking consent for data sharing, and includes a model consent form.

5. How will retention and disposal of personal data and confidential information after project completion be managed?

5.1 Retention period

State how long you plan to retain personal data/confidential information after the end of the project.

5.2 Responsible person(s)

Specify under whose authority this information will be maintained and disposed of after the project.

Guidance

Personal data should be retained as long as necessary for the specified purpose. It is acceptable under data protection law to retain personal data, used for research purposes, for long periods, subject to periodic review, if they are held for archiving purposes in the public interest, for scientific or historical research purposes, or for statistical purposes. For example, if follow-up studies are contemplated, then continued

retention is justified, providing this purpose has been notified to the individual. Care should be taken to avoid commitments to the Research Ethics Committee or participants to destroy personal data by a given time, e.g. 3 years after the completion of the project. It is better to indicate that data will not be held longer than necessary for the specified purpose(s), and schedule regular reviews of personal data holdings to determine whether they need to be retained or can be safely destroyed.

Consent forms should be held as long as required in accordance with any contractual, legal, or statutory obligations that are specific to your study; as a minimum they must be retained for as long as any related personal data/confidential information are held.

If personal data/confidential information will be retained in the long term after the completion of the research, planning should take into consideration where and under whose authority they will be held, and what provision is made for transfer of ownership should the original owners leave. It is always advisable for personal or confidential information to be owned by/accessible by more than one person, to avoid the risk of data becoming orphaned when a sole owner leaves the University. For example, a research group could maintain a personal data asset register, listing personal data held, owners of the data, storage locations, the retention schedule, and the date of next review. This register could then be updated on an annual basis and ad hoc when a review date is reached, or when any listed data owner leaves the University.

Guidance on the retention of personal data can be found in Data Protection for Researchers.