DATA PROTECTION CHECKLIST FOR RESEARCHERS

Contents

Introduction

This Checklist is for use by anyone who is planning to undertake research that will involve the collection of personal data from researchers. It consists of 9 things that you should do before you start your research.

The Checklist is a planning tool, and can be used as to inform your data management plan. A comprehensive guide to Data Protection for Researchers is available from the IMPS website at http://www.reading.ac.uk/imps-d-p-dataprotectionandresearch.aspx.

1. Specify the purpose or purposes for which you require the personal data

The purpose(s) for which personal data are required must be stated in the privacy notice supplied to research participants. If you plan to process the data for any purpose other than the proposed research, this should be clearly defined. For example, consider whether you will want to establish a database of potential participants in future research. In such a case, you would need to secure separate consent, aside from consent given for participation in the research.

2. Identify who will have access to the personal data

Access to personal data should always be on a need-to-know basis. In many cases, not all members of a research team will need to know the identities of the research participants. Participant information and linked pseudonyms can be stored in a separate online or physical location accessible by an authorised group only, e.g. the PI and Co-I, while a pseudonymised or anonymised version of the data could be made accessible to other members of the team for analysis.

If the personal data will be collected by or shared with contractors or partners outside of the University, identify these. These parties will be either Data Controllers or Data Processors.

3. Identify the Data Controllers and the Data Processors

Who determines the purposes and the manner of data processing? Who will process the personal data on behalf of the Data Controller?

If University staff or students working for or under the instruction of the University will determine the purpose and manner of processing, the University will be the Data Controller.

If project partners will process personal data under instruction from University members, they will be Data Processors.

Project partners may also determine the purposes and/or manner of personal data processing, in which case there will be more than one Data Controller.

Data Processers might be organisations conducting data collection on the University's behalf, or the providers of instruments used for data collection, e.g. online survey tools.

Any disclosure of data outside the University should take place under a contract in which the data protection responsibilities of other parties are defined. For example, where the joint Data Controller or Data Processor is a partner in a project, data protection responsibilities should be written into the research collaboration agreement. Where the Data Processor is a contractor, data protection responsibilities should be included in the service contract.

- Contact Research Contracts for advice on collaboration agreements.
- Contact IMPS or Legal Services for advice on service contracts.

4. Specify the personal data you will need to collect

Consider the nature of the information you will collect from the research participants, and in particular whether any of it, when combined with their personal information, falls into the special categories of data as defined under the GDPR. Special category data includes information about an individual's race; ethnic origin; politics; religion; trade union membership; genetics; biometrics (where used for ID purposes); health; sex life; or sexual orientation. Any personal data falling into these categories should be identified and will need to be managed with extra care.

For example, opinions of identifiable members of the public about building design are unlikely to fall within a special category of data; the mental health history of identifiable adolescents with anxiety and depression falls within the special category of data concerned with an individual's health.

Consider also what personally-identifying information you will or may collect. You should collect only the minimum necessary for your purposes. For example, for participation in an observational study, you might collect the person's name, address, email address and telephone number, to facilitate contact; for an online screening survey, you may need only to collect an email address.

Bear in mind that qualitative data collection, for example, by means of interview, observation, photography, and creative expression by participants, may yield personally-identifying references incidentally, even if interviews are conducted 'anonymously'.

5. Decide whether you will need to conduct a Data Protection Impact Assessment

You are required to carry out a Data Protection Impact Assessment (DPIA) if your processing of personal data is likely to result in a high risk to individual's interests. High risk could result from either a high probability of some harm, or a lower possibility of serious harm. Some examples of high risk data processing include processing special category data on a large scale, processing biometric or genetic

data, and processing personal data in a way which involves tracking individuals' online or offline location or behaviour.

The purpose of a DPIA is to:

- describe the nature, scope, context and purposes of the processing;
- assess necessity, proportionality and compliance measures;
- identify and assess risks to individuals; and
- identify any additional measures to mitigate those risks.

If a DPIA is required, this must be carried out with the involvement of the University's Data Protection Officer.

To find out whether you need to conduct a DPIA, you should read the University's Data Protection Impact Assessment guidance, and assess your proposed personal data processing against the DPIA Screening Checklist. The guidance and checklist can be found at http://www.reading.ac.uk/imps-d-p-dataprotectionbydesign.aspx.

Please send your completed DPIA Screening Checklist to the University's Data Protection Officer at imps@reading.ac.uk. The Data Protection Officer will then assist you in conducting a DPIA.

6. Specify the means by which the information will be collected and stored

Information is likely to be collected and stored using different media. Map your data workflow, and identify measures for securing the personal data in its holding locations and in transit between them.

Security controls such as password protection access controls can be applied to individual files, folders, storage volumes, and devices. It is advisable to use some level of access control for all digital personal data holdings, and to use encryption to protect sensitive data.

Digital channels of communication can be encrypted, so that only the sender and receiver can read the message. The University VPN provides a secure connection to the University network. If you are collecting data via online survey tools you should check that data are securely protected when transferred to and from the service provider. Email messages are not encrypted by default, and should not be used to send unprotected files containing large volumes of personal data or sensitive data. The University Encryption Policy (which can be downloaded from http://www.reading.ac.uk/imps-policies.aspx) defines requirements for the use of encryption for sending data outside the University network.

IMPS provides guidance on encryption on its web pages. IT can advise on managing access permissions to fileshares and folders on the University network.

Be especially careful when using cloud services for the storage and transfer of personal data. University OneDrive accounts are suitable for storing personal data, although care should be taken to share such data with nominated authorised recipients only.

Other cloud services (such as Dropbox, Google Drive, iCloud, etc.) are not underwritten by any institutional guarantees, and their terms of service may offer inadequate protection for personal data. For example, if data held in cloud services are stored on servers in countries outside the European Economic Area, and an equivalent level of protection for personal data is not provided in a these countries, researchers using these services to store personal data risk being in breach of data protection laws. If you are in doubt about the compliance of any cloud service with data protection laws, you should consult IMPS or IT for advice.

Here are two examples of approaches to protecting data throughout the research workflow:

Example 1

Interviews will be recorded on a password-protected audio recording device. Recordings will be transferred by encrypted VPN connection to an access-controlled folder in a fileshare managed by the PI on the University network. Once transfer is verified, recordings will be wiped from the audio device. A member of the project team will produce anonymised transcripts of the audio recordings, using their campus PC and saving all materials to the same network location.

Example 2

Screening data for an observational study, including information about an individual's health, will be collected by means of the Online surveys tool. Online surveys provides end-to-end encryption of survey responses. Data held by Online surveys will be stored on servers based in Ireland, in the European Economic Area. Data will be exported from Online surveys via encrypted connection to Excel and will be stored in the PI's University OneDrive account. Once data have been exported, they will be deleted from BOS.

For selected participants, participant information and signed consent forms will be collected on paper in private rooms at the on-campus clinic, and will be stored in a locked filing cabinet in the PI's office. Participant information will be input into the study database by the Research Assistant. Consent forms will be digitised, and stored with participant information in a restricted folder, accessible by the PI only, within the project fileshare on the University network. Paper forms will be destroyed once they have been digitised and saved. Participants will be assigned a unique identifier, which will be recorded in the restricted folder against the participant information. For the research experiments, participants will be identified by their unique identifier only. Results of laboratory-based tests will be recorded in an Excel file against the participant's unique identifier and saved to an area of the fileshare accessible to all members of the project team.

Consider also that personal data may be held in different locations during the project and for the longer term after the project.

- Contact IT for guidance on suitable University-managed services for the storage and transfer of personal data, and assistance with encryption and access controls.
- Contact IMPS or IT for advice on suitable cloud services for storage and transfer of personal data.
- Contact your School or Department administrator if you need somewhere secure to store paper-based personal data.

7. Anticipate how long the personal data will be retained

Personal data should not be retained longer than necessary, but may be retained beyond the duration of the original research project 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.

If the only purpose for which data are being collected is to undertake a given study, then personal data should not be retained longer than necessary to complete the research and publish the results. Consent forms provide an auditable record of informed participation in the research, and you should take into account any contractual, legal, or statutory obligations that are specific to your study when deciding how long to retain them; as a minimum they must be retained for as long as any personal data collected from the participants is held by the University.

Care should be taken to avoid commitments to destroy personal data by a given time, e.g. 3 years after the completion of the project. It is better to schedule regular reviews of personal data holdings to determine whether they need to be retained or can be safely destroyed.

If personal data 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 owner leave. For example, a research group could maintain a personal data asset register, listing personal data held, owners of the data, their 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.

8. Plan for pseudonymisation and anonymisation of personal data

To minimise the risk of inappropriate disclosure of personal data, it is advisable to create pseudonymised or, where possible, anonymised versions of working data. Data can often be pseudonymised for purposes of processing and analysis, with the personally-identifying information and their linked pseudonym IDs stored separately from a dataset in which individual identifiers have been replaced by the pseudonym IDs.

This may be a first step towards full anonymisation of the data. Bear in mind that the University and many funders require data supporting research findings to be preserved and made publicly available on publication of results wherever possible,[1] so it will in any case be necessary to prepare an anonymised version of the data towards the completion of research.

In the case of pseudonymised data, they will continue to be personal data as long as the identifying 'key' data is held, whether by the University or a third party. But using pseudonymisation is one way to minimise the risks of any data subjects being identified if data is shared with another person or organisation that does not have any access to the identifying data.

In one-off study scenarios it is often appropriate to destroy the linking key when the study is completed, so that data become fully anonymised. In some cases, for example where participants have been recruited for longitudinal studies, there may be a need to retain the key while anonymised versions of data are created for public release.

Bear in mind that effective anonymisation may involve much more than replacing personal names with pseudonyms, and different techniques are required for quantitative and qualitative data.

Some data collected from research participants may be difficult or impossible to anonymise effectively. Examples include: photographic images and video recordings featuring individuals, personally identifiable genetic and phenotypic data, and biometric information used for identification purposes. Where the information falls within a special category of data, extra care will need to be taken in their collection, storage and disclosure. Managed disclosure of such data in support of published results may still be possible: data archiving services such as the UK Data Service and the European Genome-phenome Archive[2] can manage controlled access to confidential data collected for research purposes. If you do plan to collect inherently confidential data, you should seek consent for controlled sharing and include in the consent form a statement such as this:

'I understand that the data collected from me in this study will be preserved, and will be made available to other authenticated researchers only if they agree to maintain the confidentiality of the information provided to them.'

[1] The University's Research Data Management Policy is available at http://www.reading.ac.uk/reas-RDMpolicies.aspx.
[2] UK Data Service: https://www.ukdataservice.ac.uk/deposit-data/preparing-data/confidential-data; European Genome-phenome Archive: https://www.ebi.ac.uk/ega/home.

- Contact the Research Data Manager for advice on planning for data sharing.

Guidance on anonymisation techniques

MRC *Good practice principles* Appendices 2 and 3 (p. 23-28).
http://www.methodologyhubs.mrc.ac.uk/files/7114/3682/3831/Datasharingguidance2015.pdf

Hrynaszkiewicz I et al. (2010), 'Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers'. British Medical Journal 340:c181. https://doi.org/10.1136/bmj.c181

UK Data Service: Anonymisation. https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation

9. Prepare the information sheet and consent form

The consent form provides documented evidence that the research participants have received the privacy notice and are aware of the lawful basis, purpose(s) and manner of data processing. Take care not to prejudice your research by poorly formulated information sheets and consent forms. In particular make sure that the purposes for which the data are being collected are clear, but do not place unnecessary restrictions on your use of the data, and notify participants if you plan to retain personal data for purposes other than those of the original research.

Make sure that the privacy notice includes the following:

- The name and contact details of the Data Controller(s) and the contact details of the Data Protection Officer;
- The purpose(s) for which the personal data are being collected;
- The lawful basis for processing the personal data;
- The categories of personal data to be collected;
- The recipients of the personal data, including third parties the data may be shared with;
- The details of the transfer of the personal data to any third countries or international organisations (if applicable);
- The retention periods for the personal data;
- The rights available to individuals in respect of their personal data, including the right to withdraw consent;
- The right to lodge a complaint, i.e. with the ICO;
- The source of the personal data (if not obtained from the individual);
- Details of the existence of automated decision-making, including profiling (if applicable).

- Contact IMPS for guidance on privacy notices.
- Contact Research Ethics or the Research Data Manager for guidance on consent forms.

Guidance on information/consent and sample information sheets and consent forms

University sample information sheet and consent form. http://www.reading.ac.uk/imps-d-p-dataprotectionandresearch.aspx

University Research Ethics Guidance Notes with sample information sheet and consent form. http://www.reading.ac.uk/internal/academic-and-governance-services/research-ethics/RECethicshomepage.aspx

UK Data Service: Consent for data sharing. https://www.ukdataservice.ac.uk/manage-data/legal-ethical/consent-data-sharing

Contacts

Team	Advises on	Email	Tel.
IMPS	Compliance with data protection law; privacy notices; encryption policy	imps@reading.ac.uk	0118 378 8981
IT	Services for storage and transfer of digital data; using encryption and access controls for data	it@reading.ac.uk	0118 378 6262
Legal Services	Service contracts with data processors	legalservices@reading.ac.uk	0118 378 6742
Research Contracts	Collaboration and partnership agreements	res@reading.ac.uk	0118 378 8628
Research Data Manager	Planning and consent for data sharing; preparing participant-derived data for public release	researchdata@reading.ac.uk	0118 378 6161
University Research Ethics Committee	Use of confidential data in research; ethical approval process; information sheets and consent forms	urec@reading.ac.uk	0118 378 7119