What to consider when you will collect personal data for your research

Awareness:

1. What is your purpose?
    - Carefully consider the purpose for which you collect the personal data. Sensitive data must only be collected and stored if you really need them to analyze your results. Please consider that any personal data that is not directly necessary for the analysis, should be clipped from the dataset as soon as possible and should not be stored anywhere.
    - Assuming you really do need that particular sensitive information, explain (especially to the data subjects) why this information is necessary for the analysis of the data. Transparency is very important.
    - Please keep in mind that you are only allowed to collect data for specified, explicit and legitimate purposes and cannot further process the data in a manner that is incompatible with those purposes;
    - The data you collect should be adequate, relevant (directly related to answering research questions) and limited to what is necessary in relation to the purposes for which they are processed ('data minimization');

2. How will you "process" the data ?
    - For instance: collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. Consider security aspects.
    - Is there a way to pseudonymize your data set and can you store the personal data separate from the collected test results? For example for age, it could be sufficient to use categories such as 20-30 years old, 30-40 years old, older than 60, etc. Alternatively, you can also use a pseudonymization key and store the files in separate and secure locations.

3. How do you save personal data and who has access?
    - Do not leave personal data unattended, store it, store it in the right folder on TU/e storage, SurfDrive or an encrypted drive with access management;
    - Do not share personal data easily with third parties;
    - Collect and store only the necessary information; clip any unnecessary information from the data set as soon as possible. See also this page on intranet about data set storage.
    - Use personal data only for the purpose for which they are intended in the first instance;
    - Use a good password and do not share it with others;
    - Lock your computer when you leave your workplace (win +L);

4. Which (confidential) personal data do you collect?
    - Does it involve confidential "personal data"? (example: Name, address, day of birth, license plate, Social security number (BSN), IP-address, racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic or biometric data, health, sex life or sexual orientation).

5. Do you forward personal data and to whom?
    - Do not share or keep the data in a form which permits identification of data subjects for any longer than is necessary for the purposes for which the personal data are processed;
    - Processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.

6. For how long do you store personal data?
    - Do not store personal data longer than necessary. Keep as less as possible as short as possible – nice to have is not an option anymore

7. What's the lawful basis of your data collection?

    The fundamental principle for handling personal data is that data must be processed lawfully and in a transparent manner. GDPR defines six lawful bases to process data. Choosing the most appropriate basis depends on the purpose of data processing. For research, the applicable lawful bases are:

    - **Consent** - When you have consent from the data subject to process their personal data. There must be a deliberate action on the part of the data subject to opt in or give consent. You must always allow your data subject the option to opt-out, unless this is impossible because you have de-identified the data set and deleted any identifiable data. See more on consent below.
    - **Public Tasks** - Carrying out tasks in the public interest
    - **Legitimate Interest** – Scientific research can be considered a genuine and legitimate reason to process data, provided the purpose does not harm the data subject's rights.

    Regardless of the lawful basis, you should inform your data subject about the fact that you collect their data, which data and for which purposes, and what will happen with their data during and after the project/experiment.

8. Transparency and consent

    Consent forms can ask participants for public sharing of anonymised research data, without specifying the purpose (this is sometimes referred to as 'open consent'). Open consent is the preferable choice for datasets which can be anonymized, or where the risk for participants' re-identification is minimal.

    Informed consent: if datasets cannot be anonymized, you have to prepare informed consent forms which explain to participants what will happen with their data during and after the project/experiment, which includes:

    - unambiguous consent of the participants including use, potential re-use, management and sharing of data;
    - strategies for safe storage of data;
    - access to own data and a clarified 'right to be forgotten' ;
    - the right to know when data has been inappropriately released/leaked/hacked

- procedures for complaints and how these will be handled
- data retention period

You must store the collected consent during and after the research for as long as you retain the data. Example of a paragraph on (open) consent, to include in the introductory text of a survey, is available at request both in English and Dutch. Please inquire at Mandy van de Sande.

Before you start your data collection:

→ Is there a low/medium privacy risk of the person involved in your research?
  - Fill out a Pre Privacy Impact Assessment (pre-PIA)

→ Is there a high privacy risk of the person involved in your research?
Large data sets of special personal data, systematic observation or monitoring of groups of people are considered high risk, as well as systematically and extensively analysing people's personal or lifestyle aspects or behaviour (for instance with the purpose to profiling).
  - Fill out a full PIA

→ The pre-PIA and PIA needs to be signed by DPO and has to include:
  - What is the goal of the data processing?
  - Which parties are involved?
  - Which data types are sensitive?
  - "Informed consents" or "public task" as "lawful basis"?
  - Which security measures are applied?

Support

→ Support in the unit USRE: Research Support Mandy van de Sande (m.v.d.sande@tue.nl)

→ Support in the Department of the Built Environment: Ms. Silvie van Dam (s.v.dam@tue.nl)

→ Expert assistance at university level is available at:
  - Privacy@tue.nl (privacy officer – Sandy Janssen)
  - rdmsupport@tue.nl (data stewards – Toine Kuiper)

→ For signature and processing of the PIA go to
  - Dataprotectionofficer@tue.nl (Data protection officer – Annuska van den Eijnden)

→ For frequently asked questions (FAQ) go to:

https://intranet.tue.nl/en/university/services/information-management-services/information-security-and-privacy/personal-data-and-privacy/faq/research/#c47991

→ First GDPR training at Built Environment (including links to presentations):

https://intranet.tue.nl/en/university/services/information-management-services/help-and-support/service-researchers/research-data-management/rdm-news/14-02-2019-1st-gdpr-training-at-built-environment/