GUIDELINE FOR DATA MANAGEMENT (COLLECTION, STORAGE AND PROCESSING OF DATA) IN STUDENT-RELATED RESEARCH

Students must comply with laws and regulations when collecting, processing and storing data as part of their studies. This applies in particular when working with personal data. Below we outline three common situations, namely:

a) New data collection;
b) Use of existing data;
c) Naturalistic observation.

We provide guidelines for the responsible use of data in all phases of data collection.

Situation A. New data collection

Step 1: request permission

Make sure you ask the participants you want to involve in the research for permission to collect the data. Active consent (consent form with signature; informed consent, see appendix) is required if it concerns the collection of (special) personal data (for definition, see appendix), video recordings, sound recordings or invasive questionnaires.

For drawing up an adequate consent form, see our instruction document [link].

Step 2: collecting data

Make sure that questionnaires are filled in anonymously, so no name, date of birth or student number on the cover page or questions about it in the (online) questionnaires; for online questionnaires: make sure that you uncheck the IP addresses when printing the answers to the questionnaire (e.g. Qualtrics has that function).

Also make sure that you do not create a contact list with email addresses of all participants in the survey in Gmail or make a call via Facebook to recruit participants.

When collecting data from the web, make sure that collecting the data that a person puts online does not violate that person's reasonable expectations with respect to his/her privacy.

In the case of video data use, as much as possible, secure devices that you can pick up from the Techsupport https://Techsupport.fss.uu.nl/ desk. And make sure you store the video data a.s.a.p. on the secure part of the faculty server (advice and access to the faculty server can be obtained via Techsupport) and remove it from the device.

For your supervisor: if the data is collected via an external body such as a school/institution/organization: check with the privacy officer of the faculty (privacy-fsw@uu.nl) whether it is necessary to enter into a processing agreement with this external body.

Step 3: processing and storing the collected data

Anonymize

Anonymizing means that you strip the data of all information that can be traced back to the person (see also appendix). Do this in consultation with your supervisor. If desired, your supervisor can ask Techsupport for further advice: https://Techsupport.fss.uu.nl/ or via privacy-fsw@uu.nl.

- Save the anonymized data a.s.a.p. on the appropriate folder on the faculty server. A folder can be created via the Techsupport staff.

- Then remove the data from the device on which they were originally stored, also remove the documents from "the trash can" of your device. Do not store the data in the cloud!

If it is not possible to anonymize the data (e.g. because you need to be able to approach participants again for follow-up research), pseudonymize.

Pseudonymize

When pseudonymizing the data, you assign a unique code to each person in the data. Next you make two data sets: one with the code and the identifying information, the so-called key, and one with the code without the identifying information, the so-called pseudonymized data (see appendix). The two datasets are stored separately from each other. Analyses take place on the pseudonymized data. Because of the unique code, the key can be linked to the pseudonymized data at a later point in time, so that you can approach the participants again if necessary. Do this pseudonymizing in consultation with your supervisor. If desired, your supervisor can ask Techsupport for further advice.

- Save the pseudonymized data a.s.a.p. on the appropriate folder on the faculty server. A folder can be created by the staff of Techsupport.

- Save the key a.s.a.p. on the secure faculty server.

- Remove the pseudonymized data and the key from the device where they were originally stored, also remove the documents from "the trash" of your device. Do not store the data in the cloud!

Linking newly collected data to existing data

Sometimes you need to link newly collected data of persons to already existing data of these persons. Think for example of linking newly collected data of students to their existing CITO scores at school. In all probability, this will be done with the help of personal data. Make sure that the use of the personal data as much as possible at the location where the data collection takes place.

It is preferable to anonymize the data after this linking, and otherwise pseudonymize it.

Step 4: data transport

If you take the data from the external organization, for example school/institution, to the university to be able to work on this further, do this:

- via a secure connection, such as Surffile sender:

https://www.surf.nl/en/surffilesender-send-large-files-securely-and-encrypted

- by borrowing a secure device from the faculty via Techsupport: https://Techsupport.fss.uu.nl/.

- by saving the data on Surfdrive: www.surfdrive.nl (and removing it from the device).

Make sure that only you and your supervisors (can) have access to the data.

- See data storage guideline: https://techsupport.fss.uu.nl/ethics-and-data-management/data-storage/

Step 5: data storage

During the process and for small files the student's own u disk. And for the final data storage and for large files YODA. See also: https://techsupport.fss.uu.nl/ethics-and-data-management/data-storage/

Project data (including field notes, recorded interviews and transcriptions, audio or image recordings and any other data) are securely stored in a password-protected, encrypted storage medium, a database to which only the researcher has access (and in the case of large research projects, other project staff as well).

Storage periods

The starting point is that data collected by the student for writing an assignment or thesis is saved. For the retention period, the periods for saving study results apply (for papers 2 years and for the thesis 7 years). If a scientific publication takes place on the basis of the data, the retention periods mentioned in the Guideline for Archiving Scientific Research for Dutch Faculties of Social and Behavioral Sciences, Version 2, July 2017 apply.

If it is an unpublished thesis it is 7 years, and if it is a published article it is 10 years.

B. USING EXISTING DATA

Student uses data at the UU (project of supervisor).

Check access to the data

- Check if and how access to the data is arranged for you.

Make sure you only have access to the data you need for your research. In concrete terms: only to anonymous data, or, if there are good reasons for pseudonymizing, the pseudonymized data.

- If you need access to the data from home, arrange this via a secure connection. For advice on this, please contact Techsupport.

Make sure you do not download data on your own laptop!

Furthermore, the following points are important:

1. If the data itself qualifies for copyright protection then the copyright lies with the student; if this is the case then a transfer agreement with the data attached must be agreed upon.

2. If there is a collection, the (composition) of the collection is also subject to copyright (without prejudice to the rights mentioned above) and these are vested in FSW.

3. In addition to the aforementioned copyrights, you also have to deal with the database right: the producer of the database is, in principle, the one who has the exclusive right to give third parties access to the database. Often the producer is the one who has made money available for the compilation of the database, including FSW.

C. NATURALISTIC OBSERVATION

Student collects data on the street, in the field, online by observing participants or other things. It is important that you do not record any personal data, nor do you make video or audio recordings. Data storage is also the same as described above for anonymous data.

Project data (including field notes, recorded interviews and transcriptions, audio or image recordings and any other data) is stored securely in a password-protected, encrypted storage medium, a database to which only the researcher (and in the case of large research projects, other project staff) has access.

Attachment:

1. Example of Informed consent form

2. Definitions of basic concepts of data management and privacy

Appendix 1:

Example information letter/template

Subject information for participation in (social) scientific research

<Title research>

<Date, Location>

Dear sir, madam,

Introduction

Through this letter we would like to ask your permission to participate in the research "<title>". The purpose of this research is....

Set up/execution of the research

Background research

<description background of the research>

What is expected of you as a participant

<description of what participating in the research means; what exactly is expected from the participant in terms of tasks, questions, time effort, duration/frequency, other forms of load>.

Possible advantages and disadvantages of the research

Compensation/Reward

Confidentiality of data processing

This research requires us to collect a number of personal data from you. We need this information to be able to answer the research question properly, or to be able to approach you for follow-up research. The personal data is stored on a different computer than the research data itself (the so-called raw data). The computer on which the personal data is stored is secured to the highest standards and only involved researchers have access to this data. The data itself is also secured by means of a security code.

Your data will be stored for at least 10 years. This is in accordance with the appropriate VSNU guidelines. You can read more information about privacy on the website of the Personal Data Authority:

https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-europese-privacywetgeving

Procedure coincidence findings

If applicable: <write something about coincidence and associated procedure; see also our FAQ>.

Voluntary participation

Participation in this study is voluntary. You can discontinue the examination at any time, without giving a reason and without any adverse consequences for you. The data collected so far will be used for the research, unless you explicitly indicate that you do not want this.

Independent contact person and complaints officer

If you have questions or comments about the study, you can contact <independent contact; Independent means: not involved in the study itself. The contact person may, in principle, be a fellow researcher (whether or not from another department) who can respond to the question or complaint. >

If you have an official complaint about the investigation, please send an email to the Complaints Officer at klachtenfunctionaris-fetcsocwet@uu.nl or contact the UU Data Protection Officer at privacy@uu.nl.

If, after reading this information letter, you decide to take part in the investigation, please sign the enclosed reply strip and hand it to the investigator(s).

Kind regards,

Name researcher(s)

Statement of consent:

I herewith declare to have read the information letter concerning research <title> and to agree to participate in the research.

Further consent use data (possibly split by type of data:)

- raw (anonymous) research data

- personal data > contact details

- special personal data (e.g. body material)

Name <Don't include any identifying information here, such as subject number or other codes, date of birth, etc.>.

Date

Appendix 2:

Definitions privacy, data management and ethics

Definition of personal data

Any information about a person that could lead to identification of that person, such as name, address, telephone number, email address, date of birth, identifiers and location information (IP addresses, Smartphone IMEI number, cookies, MAC address) and fingerprint. Identification can also occur with a combination of these data, for example, only a first name is in principle not identifying, unless it is very rare.

Definition of special personal data

Special personal data are extra sensitive data about an individual which, if processed, can have an extra negative impact on someone. Examples of special personal data are: a person's health, genetic data, biometric data, ethnicity, religion, political opinions, trade union membership, criminal record or sexual life. Special personal data only exist if they can be traced back to individuals. This is done, for example, by requesting them in combination with personal data. A combination of special personal data can also be (indirectly) identifying in a certain context.

Definition "processing" in the sense of the GPDR (to be interpreted broadly)

Any action that can be performed with the personal data falls under the legal term "processing". Examples include consulting, collecting, recording, organizing, storing, retrieving, using, distributing or destroying. For example, if students do an internship at school and they are presented with a class list, they process personal data and are expected to treat it confidentially, for example, not to copy the list and only to use it for the purpose of that moment. The same applies to (patient) files.

Lawful Basis for processing personal data

The question here is whether there is permission from the participants to process their personal data. Unit-holders can give this consent in a so-called Informed consent form (see below). If this consent is not given, it may be possible to invoke the "legitimate interest" basis or one of the other principles of the GPDR.

Informed consent

Adequately informing participants about the research, after which they can give permission for participation and use of their data.

Data minimization

Asking out only those personal data that are necessary for the purpose of the investigation. Think carefully about what personal data you ask participants. Are they all really necessary for the purpose of your research? For example: if the age of the participant is necessary to answer your research question, do not ask a date of birth but only the year or age. Do you really need the address details, e.g. to approach the participant again? Is this not possible by email?

Anonymize/pseudonymize

If personal data has been completely deleted, this will constitute anonymity of persons, and the AVG will no longer apply. As a rule, we work with anonymous data. If the data is not (yet) anonymous, we anonymize the data. Anonymizing is more than leaving out names and contact details. It is about the fact that it is not possible to identify someone with any means that can reasonably be used. The dental profession of someone in the city of Appingedam will lead to certain people with a reasonably high probability.

Sometimes it is necessary to keep personal data, because participants need to be able to be approached again (think of longitudinal research). In that case data will be pseudonymized (this is assigning a code to participants in the data, whereby the identifying information is replaced by this code (a key) and the identifying information together with the key is kept separate from the original data). By pseudonymized data we mean the research data with the code. By key we mean the personal data with the code. Pseudonymizing means that there is still personal data, because the data can still be traced, albeit more difficult, because in order to identify one must have the key and the pseudonymized data. The key is stored on a secure server.

Storage of the data

Personal data is preferably deleted, but if it is kept, it will be stored separately from the research data. See also the definition for anonymization and pseudonymization. Describe how the data is stored. Also indicate if you plan to make the data open access in the long run.

Storage period

The personal data will be destroyed as soon as they are no longer needed to carry out the investigation. The raw data will be stored for at least 10 years (and at least 15 years in the case of WMO-liable research).

Access

It is indicated who has access to the data and whether the ultimately anonymized data will be made available for open access.

In principle, unit-holders have the right to access the data for as long as their personal data are stored.

In principle, unit-holders are entitled to have their personal data deleted. This is only possible if they have not yet been destroyed.