Final theses and processing of personal data

When you are writing your final thesis and collecting data for it, you need to observe legislation concerning the processing of personal data. Pay attention to the processing of data especially if your study involves collecting personal data or you have access to databases with personal data.

The most important statute concerning the processing of personal data is the EU's General Data Protection Regulation (GDPR, EU 2016/679), which is complemented by the National Data Protection Act (1050/2018).

When you collect data for your final thesis, ***aim to collect data anonymously***. This ensures the data subjects' rights to their personal data and limits your obligations ensuing from data protection legislation.

If you receive ***access to a database containing personal data*** during your thesis work, you will be considered a ***data processor*** under data protection legislation. In such cases, you will be obligated to observe the data controller's instructions on how to process personal data.

If you are conducting ***a study that contains personal data***, you will be considered a ***data controller*** under data protection legislation and must do as follows when you process the data:

- **Consent**
  - The collection of personal data always requires ***consent*** by the data subject (see consent form). If you collect data e.g. through the Webropol system, the data subjects may give their consent e.g. by checking a box.
- **Transparent information**
  - When you collect personal data, you must always prepare a ***privacy notice*** and communicate it to the data subjects (see privacy notice template). Present the privacy notice e.g. in business enterprises when you conduct interviews or attach it to a Webropol survey.
- **Purpose limitation**
  - You may use the personal data that you collect only for the purpose indicated in the privacy notice.
- **Data minimisation**
  - You may not collect personal data that you do not actually need for your final thesis. According to the GRPR, the personal data must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.
- **Accuracy**
  - Make sure that the personal data you collect is accurate. Incorrect and erroneous personal data must be erased or corrected without delay.
- **Storage limitation**
  - Destroy all personal data when you no longer need it for the preparation of your final thesis. You must observe the storage periods mentioned in the

privacy notice. If the entire collected database cannot be destroyed with consideration to the purpose of use, the data must be made anonymous so that the data subjects can no longer be identified from the research data.

- **Integrity and confidentiality**
  - Make sure that the personal data is processed and stored securely. The privacy notice must also mention how the collected data is safeguarded. Store manual data in locked facilities and make sure that only persons who need to process it for the preparation of your final thesis have access to it. Likewise, if the data is in electronic form, you must make sure that only persons who need to process it for the preparation of your final thesis have access to it.
- **Rights of the data subject**
  - Data subjects have the rights to their personal data. The privacy notice informs data subjects of their rights regarding its processing. The data controller must comply with a data subject's requirement e.g. to correct data concerning the data subject.

Terminology of the General Data Processing Regulation

Personal data

The definition of personal data is rather extensive because personal data refers to any information concerning an identified or identifiable natural person (data subject). An identifiable natural person is someone who can be identified directly or indirectly especially based on identifiers such as a name, an identification number, location data, an online identifier, or one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. *Therefore, the definition of personal data is ambiguous. If you have difficulties determining whether a piece of data is personal data, it is always safer to treat it as personal data.*

Special categories of personal data

The processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation is prohibited. This does not apply, however, if the data subject has given explicit consent to the processing of the data or if processing is necessary for other reasons specified in article 9 of the GDPR.

Genetic data means personal data relating to the inherited or acquired genetic characteristics of a natural person, giving unique information about the physiology or the health of that natural person and resulting, in particular, from an analysis of a biological sample from the natural person in question.

Biometric data means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, such as facial images or fingerprints, which allow or confirm the unique identification of that natural person.

Data concerning health means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveals information about his or her health status.

Personal data filing system

A personal data filing system may be any collection of personal data where data is available under certain conditions.

Data controller

A data controller means a natural or legal person, authority, agency or other body that alone or together with another defines the purposes and means of personal data processing.

Data subject’s consent

Consent refers to any clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data. Data subjects should preferably give their consent in writing because it enables the data controller to prove that the data subject has agreed to the processing of personal data.

Anonymisation

Anonymisation refers to the processing of personal data in a way that makes the data subject unidentifiable. The data can be transformed into statistical form or aggregated so that data on individual persons can no longer be identified. For data to be anonymous, identification must be irreversibly impossible in a manner that the data controller or a third party can no longer make the data identifiable. Simply removing names and other identifiers may not make the data anonymous. The collected data may contain detailed information (e.g. a professional position or rare disease) that enables indirect identification. *Anonymised data is not considered personal data and is not governed by data protection legislation.*

Additional information on data protection, the obligations of data controllers and personal data processors, and the rights of data subjects are available e.g. on the website of the data protection ombudsman: https://tietosuoja.fi/etusivu.

The EU's General Data Protection Regulation is available here: https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=celex%3A32016R0679.