

Application to the Ethics Committee

Research involving human participants

Guidance on the distinction between “anonymous/anonymised research data” and “personal data”

Introduction

This document is not intended to provide a detailed or comprehensive treatment of the subject but serves as guidance for applications to the Ethics Committee, helping to clearly distinguish between different types of data in the context of data protection (see point 6 of the [checklist for research involving human participants](#)). Therefore, only the most common issues in research projects at BOKU University are addressed.

The term ‘research data’ includes data from online surveys, transcripts of interviews, texts, samples, measurement data, etc., which are created, developed, or evaluated in scientific work. Research data can be either anonymous or personal.

What is anonymous or anonymised research data?

In the context of an application to the Ethics Committee, this refers to anonymous¹, anonymised² or pseudonymised³ research data (in other words: it is no longer possible to identify the natural person from whom the data originated). Therefore, it is not subject to GDPR and does not need to be protected from unauthorized access. In the context of [Open Research Data](#), publication in repositories (e.g., [Zenodo](#)) is even desirable. While anonymous, anonymised or pseudonymised data does not need to be protected, secure storage to prevent data loss is of course advisable.

What is personal data under GDPR?

Personal data under the [General Data Protection Regulation \(GDPR\)](#) includes any information related to an identified or identifiable living natural person, i.e. a person who can be identified, directly or indirectly, e.g. by reference to a name, a location, or an online identifier (see [Art. 4 \(1\) GDPR](#)). Several items of partial information that, together, can lead to the identification of a natural person, are also considered personal data.

The GDPR also regulates the processing of "special categories of personal data" (see [Art. 9 GDPR](#)), which are particularly sensitive and require special protection. This includes genetic and biometric data, among others.

¹ In the case of anonymous data, it was already ensured at the time of data collection that the identification of a natural person is impossible.

² In the process of anonymisation, the information that would enable the identification of natural persons is removed from the research data.

³ Pseudonymisation is a term from the GDPR (see [Art. 4 \(5\)](#)), and refers to processing personal data in such a way that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Differences concerning data protection and retention period

	Anonymous / anonymised / pseudonymised research data	Personal data
Subject to GDPR?	No	Yes
Protection from unauthorized access?	Not required. The data can be made available to other researchers.	Necessary and essential to protect the rights of natural persons. The data should only be accessible to the researchers who process it within the scope of the project.
Retention period	Data should be retained for at least 10 years in accordance with Good Scientific Practice (see § 2f para. 3 FOG).	Data must be erased once the purpose for which it was collected and processed has been fulfilled (see Art. 5 para. 1 lit. e GDPR).

Why collect and process personal data?

In many cases, personal data is required, for example, to arrange interview appointments or to invite interested parties to workshops. Email addresses may also be needed towards the end of the project to inform participants about the results. Addresses may also be necessary to analyse travel routes.

This purpose must be explicitly stated in the application to the Ethics Committee, as it determines the duration for which this personal data is retained. Indeed, personal data should only be stored as long as necessary for the purposes for which they are processed. In most cases, they are to be erased at the end of the project (or a few months thereafter), as the purpose for which they were processed has been fulfilled.

Specific examples

Whether the collected data constitutes personal data under GDPR must be assessed on a case-by-case basis. The following examples from typical research projects at BOKU serve as guidance.

Examples of personal data under GDPR:

- Name, telephone number, address, personal email addresses
- IP address collected during online surveys, as it also allows to identify natural persons
- Audio files from interviews, as people can be recognized by their voices
- Video files and photos where individuals are identifiable
- Even several pieces of partial information can collectively lead to the identification of a natural person, e.g., postal code, street names without house numbers in sparsely populated areas, GPS coordinates of fields or buildings, farm size and type. This means that even if no name, address, or email address is collected in an online questionnaire aimed at farmers, but detailed information about their farm is gathered, it is likely not an anonymous data collection.⁴

Examples of anonymous/anonymised/pseudonymised research data:

- Online surveys in which the invitation to participate is sent by a market research institute, and no IP addresses are collected. In most cases, these can be classified as anonymous.
- If data such as gender, profession, age or postal code is collected in a survey, this is usually not personal data in accordance with the GDPR, as in most cases, a natural

⁴ It may therefore be useful to consider whether all data is really necessary for the analysis, or if some can be omitted to enable anonymous data collection.

person cannot be identified with this data. However, this depends on the context: If a study explicitly lists professions practiced by very few people, it may be possible to infer a natural person (e.g., “gender: female” and “profession: rector”).

- Interview transcripts where all names and other personal or personally identifiable information (e.g., proper names of people, companies, locations) have been removed during transcription.⁵
- Text files analysing videos that do not allow any conclusions about the filmed person.

For further information on applications to the Ethics Committee, see: <https://short.boku.ac.at/ethics>

⁵ It should be noted that in interviews, complete anonymity often cannot be guaranteed, as neighbours, colleagues, family members or others may be able to make assumptions about the identity based on the statements. Participants must therefore be informed of this in advance, even if, in the context of the application to the Ethics Committee, these anonymised transcripts are considered "anonymised research data".