

CEBES – Data storage

Below, you find aspects to consider when storing survey data; in particular if they have been recorded using online media (Internet, smartphone, etc.). This guideline is part of step 3 of the CEBES procedure (source: ordinance of the Human Research Act, Art. 5, 25 and 26; “Richtlinien für Integrität in der Forschung und gute wissenschaftliche Praxis an der ETH Zürich Art. 11/12, „Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten der Deutschen Forschungsgemeinschaft; explanatory note „Informationsverwaltung“ of the Data Protection Authority of the Canton of Zurich).

The following rules refer to storing primary survey data:

- 1) The primary data should be stored on paper or (in most cases) in digital form by the principal investigator (PI) of the study on a suitable storage device. The PI restricts the access to the data to those persons that need this access for fulfilling their study-related tasks. In addition, the PI determines which study personnel will still have access to the data after they left the project team and for what purpose they can use this data.
- 2) Every set of primary data should also include the relevant meta-data: on the one hand bibliographic facts (name of the PI, name of the data set, place and time of publishing the data, technical data, e.g. related to the format), and on the other hand content-related meta-data that describe the data set. This includes information on how the data was gathered and the question that motivated the study. In addition, all procedural steps that have been undertaken with the primary data (statistical analyses, transformations etc.) have to be documented such that the published findings can be reproduced completely.
- 3) The primary data should be stored in an anonymous form (exception, see point 4). This means that data items that allow for an identification of the individual person without disproportionate efforts should be deleted or made unrecognizable. This includes in particular name, address, date of birth, IP-address and clear identifiers like phone numbers, license numbers or social security number. In some cases, also indirect identifiers like body height, hair color, profession, functions etc. may have to be deleted.
- 4) If the primary data includes data sets that have to be attributed to individual persons or if the study included a right to withdraw the information given, then the PI has to code the data. The code has to be stored safely by a person (determined by the PI) that is not part of the study team. In some cases, one may also choose the method of pseudonymization, i.e. the study participant chooses a pseudonym that only the study participant himself/herself knows.
- 5) After completion of the study, the primary data has to be stored for at least five years. The PI has to take measures to prevent unauthorized access, accidental disclosure, manipulation, copying or deletion of the primary data. If the access to data is restricted, this has to be indicated on the data set (“confidential”). Primary data that has been surveyed in a study that did not need mandatory authorization by a Cantonal Ethics Commission should be offered to the “Staatsarchiv” for archiving. They will decide, whether the data should be stored further or not (Archivverordnung §10; see also point 6). If the data consists of health-related (non-

anonymized) personal data that has been surveyed in a study that needed mandatory authorization, then special conditions apply regarding the duration of the storage.

- 6) Researchers are not obliged to make primary data available to third parties outside of the study team before they have analyzed and published the data. Reserved is the right to access for commissions (e.g. ethics commissions). After data analysis and publication of the results of the study, the data may be stored in a repository. In this case, anonymization is mandatory. In addition, meta-data as described in point 2 has to be stored as well. The consent form should include the remark that the data will be stored in a repository.
- 7) Primary data that has been surveyed in a study conducted at the University of Zurich is property of the University unless other agreements have been made with external project partners.