EUI EUROPEAN UNIVERSITY INSTITUTE

GUIDE ON

GOOD DATA PROTECTION PRACTICE IN RESEARCH

With the support of the
Erasmus+ Programme
of the European Union

The European Commission supports the EUI through the European Union budget. This publication reflects the views only of the author(s), and the Commission cannot be held responsible for any use which may be made of the information contained therein.

CONTENTS

SCOPE AND PURPOSE OF THE GUIDE

The present Guide provides an overview of the major concepts of data protection and privacy in research. It aims at raising awareness for these concepts amongst the EUI's academic community and at assisting researchers with the preparation of their project proposals or submissions to the EUI Ethics Committee.

The third edition reflects the developments in the EUI's Data Protection Policy since 2017. The main novelty is President's Decision No. 10 of 18 February 2019 regarding data protection at the EUI. President's Decision 10/2019 has adapted the EUI's Data Protection Policy to the new General Data Protection Regulation (GDPR) and Regulation 1725/2018.

The goal of this guide is also to provide researchers with a handy tool to guide them through the daily work on their research project.

Please study carefully the relevant notions and measures. You should send requests for review by the Ethics Committee early, especially in case funding authorities (such as the European Research Council) have identified issues in relation to personal data. The researcher should also fill in the sample Notification Form (annexed to this guide). Likewise, you should provide detailed information on the procedures for data collection, storage, protection, retention and destruction. In general, the researchers must make sure that they will comply with EUI as well as national and EU rules.

The information contained in the Guide is of a general nature only. It is not intended to address the specific circumstances of any particular individual or project but rather informs about the main aspects of data protection in the context of research carried out at the EUI. It does not provide binding legal advice.

This Guide should be consulted in parallel with the EUI's Data Protection Policy as well as with the EUI Code of Ethics in Academic Research.

Annalisa Baldassarri provided invaluable help for this revision.

The document remains open to adaptation, invention and modernisation whenever legal, pragmatic or technological developments make room for it. Comments and suggestions for improvement are welcome at **data_protection_officer@eui.eu**.

Dr. Günter Wilms
Data Protection Officer (DPO)
European University Institute
April 2019

1. INTRODUCTION

Privacy and data protection are fundamental rights, which need to be protected.

Privacy **can mean different things in different contexts and cultures**. It is therefore important to detail the purpose of the research according to the different understandings of privacy. For example, in 'covert research' researchers should take into account the meanings of public and private in the contexts they are studying. Covert observation should only proceed if researchers can demonstrate clear benefits of the research, when no other research approach seems possible and when it is reasonably certain that no one will be harmed or suffer as a result of the observation.

Data protection aims at guaranteeing the individual's right to privacy. It refers to the technical and legal framework designed to ensure that personal data are safe from unforeseen, unintended or malevolent use. Data protection therefore includes e.g. measures concerning collection, access to data, communication and conservation of data. In addition, a data protection strategy can also include measures to assure the accuracy of the data.

In the context of research, privacy issues arise whenever data relating to persons are collected and stored, in digital form or otherwise. **The main challenge for research is to use and share the data, and at the same time protect personal privacy**.

In order to ensure respect for data protection and privacy, the EUI has adopted a Data Protection Policy[4] that must be respected by all EUI members and which is inspired by the EU data protection rules.

As a source of further reference, the EU General Data Protection Regulation (GDPR) contains a number of key principles for the processing of personal data[5]. This Regulation provides the legislative framework for data protection and privacy issues in the Member States of the European Union. In the same way, EU Regulation 1725/2018 provides the rules for the processing

4 President's Decision No. 10 of 18 February 2019 (full text available here) revising the EUI's Data Protection Policy

5 The GDPR is directly applicable in the Member States since 25 May 2018 (full text available here)

of personal data by the EU institutions, bodies, offices and agencies[6]. When the planned research includes processing of data carried out in a EU-Member State, applicants need to identify the applicable local or national legal requirements and the competent authorities, which can provide any necessary authorisations.

2. PERSONAL DATA

Who is the person behind the information?

In legal terms, '**personal data**' means:
any information relating to **identified** or **identifiable natural persons** referred to as 'data subjects'. '**Identifiable persons**' can be identified:

- **directly**, or
- **indirectly**, in particular by reference to an identification number or to one or more factors specific to their physical, psychological, genetic, mental, economic, cultural or social identity.

A. '**Any information**': the term calls for a wide interpretation. The concept of 'personal data' includes any sort of information about a person. It covers 'objective' information such as the age of a data subject, and 'subjective' information such as opinions or assessments.
B. '**Directly**' **identified or identifiable persons**: this notion includes the name of the person as the most common identifier.
C. '**Indirectly**' **identified or identifiable persons**: this notion typically relates to the phenomenon of 'unique or rare combinations' of more identifiers. There are cases, in which prima facie the extent of the available identifiers does not allow anyone to single out a particular person, but that person might actually be 'identifiable', because that information combined with other pieces of information may still allow distinguishing that individual from others. Therefore, to establish whether a person is identifiable, you will need to check whether the combination of pieces of information, which are available to the Data Controller or to any other person, can lead to the identification of the person concerned.

6 Regulation (EU) 1725/2018 entered into force on 18 December 2018 (full text available here)

Some characteristics are so unique that someone can be identified with no effort ('present Prime Minister of Spain'), but a combination of details on categorical level (age category, regional origin, etc.) may also allow the identification of a natural person, in particular, if combined with other information on the person.

The identification by name is the most common situation in practice, but a name itself may not be necessary in any case to identify the individual. Other 'identifiers' or any combination of these may be used to identify a person.

Examples of possible identifiers: physical characteristics, pseudonyms, occupation, address etc.

However, if the identification of a data subject involves an excessive effort, the individual is not considered 'identifiable' (for 'anonymisation' see also below section 8, p. 16).

Regarding the format or the medium containing the relevant information, the concept of personal data includes information available in whatever form such as alphabetical, numerical, graphical, photographical, acoustic, etc. It includes **information kept on paper as well a s any information stored digitally or on a videotape**, for instance. In particular, **sound and image data may represent information on an individual and therefore can be personal data**.

An e-mail will for example contain 'personal data'!

Previously collected data: 'secondary processing'

Secondary processing takes place when you process data that were originally collected for another purpose (e.g. in the context of a different research project) but are now being processed without the knowledge or new consent of the data subjects. The further processing of the data is lawful only as far as the new processing is compatible and covered either by the data subject's initial consent or a separate, new consent.

TIPS:

- provide details of the initial data collection as well as of the informed consent procedure

- if you use public available data, provide details of the source/s and ensure that your intended use of the data complies with any terms and conditions established by the Data Controller.

Awareness Q#1: Will you collect, store, or in any other way process any type of personal data within the framework of the research?

1.1 Categories of data processed:

- personal data (e.g. name, home address, e-mail address, location data, etc.)
- sensitive data (e.g. religious believes, political opinions, medical data, etc.)
- previously collected data.

Awareness Q#2: What categories of human participants/data subjects will your research involve?

2.1 Indicative categories of human participants:

- patients
- volunteers (for surveys, health research, etc.)
- workers (e.g. research lab personnel)
- participating researchers 'list
- children or teenagers under legal age
- vulnerable adults
- others…special population groups? Population of developing countries? Etc.

What is 'processing'?

In legal terms, 'processing of personal data' means: '**any operation or set of operations which is performed upon personal data**, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure, transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction'.

In a nutshell, anything you do with personal data is considered as processing. Here are a few examples:

- you create a mailing list or a list of participants
- you manage a database
- you share data with a third party.

Awareness Q#3: Are you familiar with the principles for processing personal data?

These principles are set out in the EUI's Data Protection Policy (President's Decision No. 10/2019 of 18 February 2019)complemented when necessary by privacy and data protection laws of the countries where data are processed (e.g. collected or stored).

You can find further explanations in guidance documents and in model operational documents on how to process personal data (some examples in the annex I to this guide, p. 22).

Awareness Q#4: Which is the applicable legal framework in case you process the data in more than one jurisdiction?

When you process data in more than one country, the Data Controller (researcher, principal investigator) must provide detailed information regarding the applicable legal framework in the countries where the processing of data is going to take place.

As a matter of law and principle, t h e Data Controller (researcher, principal investigator) must comply with data protection legislation in the country/ies where the research will be carried out.

The EUI is an international organisation. Therefore, if the research is exclusively carried out at the EUI's premises, the applicable data protection framework is the EUI's Data Protection Policy, complemented when necessary by local privacy and data protection laws.

For research and data collection carried out under a different jurisdiction (EU Member States, non-EU countries or in the frame of any other international organisation which has its own Data Protection Regulation/ Policy), researchers should make sure they comply with the respective data protection and privacy requirements (including prior-authorisations and notification requirements to National Data Protection Authorities/local Data Protection Committees).

3. BASIC PRINCIPLES FOR PROCESSING PERSONAL DATA

A healthy diet (necessity & proportionality)

Whenever you process personal data, you must keep in mind that the processing must be necessary and proportionate in relation to:

a. What?
b. Why?
c. How?
d. For how long?

'Data quality', as one of the conditions, imposes that the data are processed:

- for specified, explicit and legitimate purposes and not further processed in a way incompatible with these purposes
- only when adequate, relevant and not excessive in relation to the purpose/s (e.g. by minimising collected information/database fields)
- fairly and lawfully
- accurately and kept up to date
- in line with data subjects' rights, including the right to be forgotten
- in a secure manner
- in a form, which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed
- for no longer than necessary for the purposes for which they were collected ('retention period')
- under the responsibility and liability of the Data Controller, who ensures and demonstrates for each processing operation compliance with the Data Protection Policy.

Tips to ensure compliance with necessity and proportionality principles:

- design proper data retention and deletion plans already at proposal stage
- consider automated deletion of certain types of data during the carrying out of research and introduce a data storage scheme for data kept after the project is completed.

Moreover, you need to take adequate precautions when you transfer personal data to third parties to fulfil 'data quality' (see also below '7. Data Transfer', p. 14)

The length of time for keeping the data depends on the purpose for which the data were collected or further processed. Therefore, once the data are no longer needed to fulfil the purpose of their processing, they should either be deleted or kept in anonymous form if they serve historical, statistical or scientific uses.

N.B.: Personal data collected by the EUI for research purposes can be further processed only for the scientific objectives for which they were collected (Art. 19 of the EUI's Data Protection Policy, 'secondary processing').

4. PROCESSING OF SENSITIVE DATA

Fragile – Handle with care!

Article 2 of the EUI's Data Protection Policy indicates some categories of data that are more sensitive than other personal data and therefore require special treatment ('sensitive data'). **Sensitive data are those revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, genetic data, biometric data, data concerning health and data relating to sexual orientation or activity**.

Examples of sensitive data:

- membership in a religious or political group
- sexual orientation
- health-related records (e.g. patient records, biometric data, medical photographs, diet information, hospital information records, biological traits and genetic material)
- criminal records or information about law investigations
- localisation data such as visas, residence, GPS satellite localization recordings or other geographic recordings.

As a rule, the processing of sensitive data is prohibited. However, Article 8 of the EUI's Data Protection Policy provides for specific circumstances, which allow for the processing of sensitive data. The most common in research is upon the **data subject's *explicit* consent**.

If you intend to process sensitive data in the course of your research, or if there is a possibility that sensitive data may be processed (unintended processing of sensitive data), you will need to satisfy stricter conditions for processing the data lawfully. You may need to provide more solid justification to the Ethics Committee.

Awareness Q#5: Are all the sensitive data, which you plan to collect, necessary and relevant for the research question?

You will need to explain the reasons behind the proposed data collection (always perform a proportionality & necessity test): you should amalgamate data from different sources only when you are sure that this is legally possible.

5. INFORMED CONSENT

May I?

Informed and free consent: the most pivotal principle for legitimate processing. A comprehensive informed consent is a crucial requirement in research.

Appropriate use of consent: consent is a weak basis for justifying the processing of personal data since data subjects have the right to withdraw their consent at any time and without justification and have their data deleted, as consequence.

The use of consent 'in the right context' is crucial.

Absence or weakness of the elements for valid consent creates vulnerability and, in practice, weakens the position of data subjects. The lack of valid consent can lead to legal challenges and liability claims against the researcher.

Definition in the EUI's Data Protection Policy:
'the data subject's consent means any freely given, specific, informed and unambiguous indication of agreement to personal data being processed'

Main aspects of Informed Consent:

1. '… **freely given** …' consent can only be valid if data subjects are able to exercise a real choice. Significant negative consequences, such as deception, intimidation or coercion if they do not consent undermine the validity of consent.

In that regard, the potential participants (data subjects) must be given sufficient information in order to be able to make an unrestricted choice of whether or not to participate. The participants 'choice shall be based on an understanding of the risks as well as of the alternatives, in an environment which is free from any coercion.

2. '… **specific** …' to be specific, consent must be intelligible: it should refer clearly and precisely to the purpose/s and the consequences of the data processing. It cannot apply to an open-ended set of processing activities. This means in other words that the context in which consent applies is limited and that blanket consent without any specification of the exact purpose/s of the processing is not valid.

Consent must be given in relation to the clearly identified aspects of the processing. These aspects include notably the categories of data processed and the purpose/s of the processing.

Consent refers to reasonable processing, which is proportionate and necessary in relation to the purpose/s. It is generally sufficient to obtain consent only once for different operations as long as they are covered by the data subject's reasonable expectations (informed consent).

3. '… **informed** …' prior information (appreciation and understanding of the facts and implications) is a precondition for valid consent. Data subjects must be informed about their right to withdraw the consent at any time and without any justification and to have their data deleted, as consequence.

The data subjects concerned must be given, in a **clear and understandable manner, accurate and full information** of all relevant issues such as:

- nature of the data processed
- purpose/s of the processing ('secondary processing', if applicable)
- recipients of possible transfers
- the time-limits for storing the data
- rights of the data subject (Article 16 of EUI's Data Protection Policy);
- absence of negative consequences if consent is not given.

Quality, Accessibility and Transparency of information are key requirements!

4. The Data Controller must be able to demonstrate that the data subjects have consented to the processing of their personal data. Participants need to agree that their data will be used for a specific research purpose and must be aware of the meaning of such use.

Examples of important aspects that researchers must take into account:

- the power relationship (for instance: factual or legal hierarchy, economic dependency)
- between researcher and research participants
- the vulnerability of the population under study
- the impact of the research results on individuals and communities (no stigmatization or discrimination).

You should thoroughly consider all relevant aspects from this list in your research proposal.

When writing a research proposal you should make sure that:

a. you show a detailed understanding of the nature of the information which you have to provide to potential participants

b. you provide the necessary information by using plain and clear language, so that it will be understandable to potential participants. The information provided must allow the participants to make a decision of whether to take part in the research based on free will– i.e. participants' decision not to take part in a survey, should not lead to any negative consequences or the impression thereof

   TIP:
   the most convenient way to ensure and demonstrate this is to produce a draft information sheet and attach the informed consent protocol (use the EUI's model consent form, annex II to this guide, p. 27) to the application.

c. pay close attention to the way research participants are approached
   Caution: In case consent is obtained only from family or community leaders ('gatekeepers') to approach individuals, this should not substitute an individually obtained consent. In case of people not able to give a valid consent (e.g. children), it will be necessary to obtain the consent from the parents or legal representatives and children`s assent.

   If research consists of fieldwork, obtaining informed consent might be an ongoing process, rather than a onetime event. Meaning that, in such a case, the process for obtaining the informed consent might evolve differently from what was anticipated before starting the research.

   Consent should be renegotiated if the inquiry moves in an unanticipated new direction. The methodological limitations of gaining

'fully informed' consent has to be made clear at the outset.

d. when seeking to obtain individual written consent from research participants, take into account the cultural and ethical norms of the population(s) under study.

Tips to address special circumstances:

- in case a written consent does not respond to the ethical norms of those studied, provide alternative ways of obtaining consent (such as recording the oral consent, the presence of witnesses, all procedures used must be documented)
- in case of participants who for any reason are not fully capable of understanding and expressing their will, replace the informed consent by other equivalent measures
- in case of observational studies, obtain informed consent from all participants and approval from the gatekeepers before the beginning of the study. If individuals cannot be identified during the observational studies individual consent should be sought after the study is finished
- observation of people in a completely public environment might not require consent. In this case, researchers would have to demonstrate that their study in no way alters the usual behaviour of the people under scrutiny and that their privacy is respected.

What about children and vulnerable adults?

If you are involving children in your research **you should devise appropriate strategies of informing them** about their participation (for example by using audio or video materials, posters, flyers, suitable to their age and understanding). Children who are capable of forming their own views should be granted the right to express their views freely in all matters affecting them, commensurate with their age and maturity.

Awareness Q#6: Do you have the necessary legal permission(s) to process the data?

- If you gather data directly from individual participants, is the planned procedure to obtain their informed consent effective?
- If personal data have been collected in the frame of previous research projects (i.e. secondary processing), you are still required to document informed consent for your research project. If you process a data set, previously gathered either by you or by another person for a different

purpose, you shall be able to answer the following questions:

- does the initial informed consent cover this further processing of the data, or
- do I need to obtain a completely new informed consent for the proposed study?

These options can be discussed along with the EUI's Ethics Committee and/or the Data Protection Officer or other competent body.

6. DATA SECURITY

Under lock and key

To process data in a secure manner you must:

- take technical and organisational measures to prevent any unauthorised access
- establish clear access rules
- organise the processing in a way that gives you the best possible control, for example by allowing for tracking of access (logbook)
- if someone processes the data on your behalf, make sure that this processor ensures for appropriate security safeguards.

These requirements are set forth and further outlined in Articles 11 and 22 of the EUI's Data Protection Policy.

In practical terms, these measures could result in:

6.1 User authentication

The way to verify the identity of a user:

- one-factor: 'something a user knows', e.g. a strong password.

  Key aspects of a strong password:
  - length (the longer the better)
  - a mix of letters (upper and lower cases), numbers and symbols
  - with no ties to your personal information, and no dictionary words.

- two-factors: 'something a user has', e.g. a signed digital certificate in a smart card.

6.2 Access control

A mechanism to allow or deny access to certain data:

- based on predefined user lists and access rights, e.g. who can access what and access permissions (read, modify, etc.)
- based on the functions of each user within the project
- role based – attribute based

6.3 Storage security

Storing data in a way that prevent unauthorised access, for example by:

- operating system controls (authentication & access control)

- use of passwords to access electronic files (e.g. use the text editor function to save a document password-protected)
- local encrypted storage (enable the full disk encryption, enable the file system, enable the text editor encryption)
- database encryption: turning data into a form that makes them unintelligible (for anyone not having access to the key)

N.B.: Your storage concerns are equally important even when your data are on your local PC, your portable storage device or in the cloud storage!

6.3.1 Recommendations for the use of cloud computing services

The use of cloud computing services (CSP) should not lower the level of protection of personal data.

TOOLS:

EUI Data Controllers (researcher/principal investigator) must ensure that the following requirements are met:

- CSP should provide sufficient guarantees for data protection:
  - EUI is sole controller, CSP may process data only on the basis of EUI's instructions
  - CSP must respect EUI's data protection policy or EU-rules (e.g. promptly respond to requests for access, blocking, rectification and deletion)
  - processing and any sub-processing, should take place within the EU
  - a legally binding contract must stipulate those conditions and provide for enforceable sanctions in case of violations.
- EUI Data Controller (researcher/principal investigator) must actively monitor the implementation of the required safeguards and other contractual provisions.

6.4 Communication security

Safe electronic communication for transferring the data can take the following forms:

- encrypted communication (SSL/TLS); (e.g. use web services whose URL starts with 'https://' and not only http://)
- firewall systems and access control lists (e.g. make sure the firewall service is enabled on your PC)

- anti-virus & anti-malware systems
- protect data and data carriers when they are physically transferred (paper notes, laptop etc.).

6.5 Other IT technical controls

- Back-ups: necessary for the availability of the systems and information
- PC configuration: security-aware settings at user level (e.g. installing security updates, anti-virus protection, local back-ups, blocking of certain software installation, etc.).

Awareness Q#7: How will you ensure secure access to the data processed?

Secure access policy has to be clearly defined and must be revised periodically. It needs to be proportional to the risks involved and the sensitivity of the data. In other words the more sensitive the data or the more vulnerable the participants or the researchers themselves, the higher the standards for security. I t must clearly state the type of safeguards that will be implemented (password protection, encryption, 'need to know basis' principles.

Awareness Q#8: How will you ensure secure storage of the data?

- data structure and format:
  you might need to compartmentalise data storage, such as databases. For example, keep sensitive data strictly separated from other personal data and encrypt the databases. You should also define how to treat incidental findings, i.e. information you come across more or less by coincidence.
- location and hardware:
  conservation methods need to be specified. Preferably, a non-WAN connected computer server or HARD disk should be used. Data should not be stored on a memory stick or other media easy to lose or access.

7. DATA TRANSFER

Sharing while caring

'Data transfer' would normally imply at least the following elements:

- the communication, disclosure or otherwise making available of personal data from the researcher to a third party regardless of the medium, **including but not limited to movement across a network, physical transfers, transfers from one media or device to another, or by remote access to the data**
- conducted with the knowledge or intention of the researcher that the third party will have access to it.

The concept includes '**deliberate transfers**' and '**permitted access**' to third parties.

Transborder flows of personal data means the movement of personal data across national borders by any means, including access to data from outside the country where they were collected as well as use of cloud technologies for data.

Article 17 of the EUI's Data Protection Policy allows transfers of personal data to a third party **only when all parties of the transfer have in place adequate safeguards for the protection of personal data**. Transfers are allowed under the following conditions:

- as long as the data are necessary for the legitimate performance of tasks covered by the competence of the recipient, or
- if the recipient establishes that the data are necessary for the performance of a task carried out in the public interest or subject to the exercise of public authority, or
- if the recipient establishes the necessity of having the data transferred and if there is no reason to assume that the data subject's legitimate interests might be prejudiced.

It is the responsibility of the Data Controller (principal investigator/ researcher) to conduct a specific adequacy assessment of the data protection system of the recipient.

FOR INTRA-EU/EEA TRANSFERS

Third parties within the EU shall have in place adequate safeguards for the protection of privacy compatible with the applicable EU Data Protection legislation (i.e. GDPR and Regulation 1725/2018).

FOR TRANSFERS OUTSIDE THE EU/EEA

[N.B.: After BREXIT, also the United Kingdom of Great Britain and Northern Ireland!]

Special precautions need to be taken when personal data are transferred to third parties outside the EU/EEA that do not provide f o r EU-standard data protection.

Without such precautions, the high standards of data protection established by the EUI's Data Protection Policy and those applied at EU level would quickly be undermined, given the ease with which data can be moved around in international networks.

TIP:

The European Commission has the power to determine whether a country outside the EU offers an adequate level of data protection. It publishes a list of its adequacy decisions on its website (here: EC's adequacy decisions).

DATA TRANSFER TO THE US

Data transfer to the US deserves special attention.

It is doubtful whether the US provides a standard of data protection comparable to the one in the EU. The so-called 'Safe Harbour' decision was invalidated by the European Court of Justice[4]. Also the new framework for transatlantic data flows, the EU-US Privacy Shield[5], received critical comments from data protection authorities[6], more recent developments put those achievements into question. The Executive Order on Enhancing Public Safety in the Interior of the United States waters down the protection granted by expressly excluding 'persons who are not United States citizens

4 Judgment of 6 October 2015 in Case C-362/14, Schrems v Data Protection Commissioner, press releases available under: https://publications.europa.eu/s/k0BT

5 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.207.01.0001.01.ENG&toc=OJ:L:2016:207:FULL

6 https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm

or lawful permanent residents' from privacy policies[7]. This order risks undermining the protection granted by the 'EU-US Privacy Shield'.

For those reasons, it is necessary to carefully evaluate on a case-by-case basis whether personal data can be transferred to the US.

If you will transfer personal data from one jurisdiction to another, make sure that data protection requirements are met in both the origin and the destination jurisdictions.

Research participants must know whether their data might be transferred to third countries. This must be either explained verbally (risk for Data Controller: more difficult to document compliance) or, preferably, provided in some written format that research participants have agreed to the transfer – i.e. via their consent, which is recorded as evidence.

TIPS CONCERNING DATA TRANSFER:

- if data processing is outsourced, remove personal data, where practicable and as far as it is possible, so that only a pseudonymous ID number is used to link individual-level data with participants' identities
- assess the level of protection afforded by a third country or international organisation in the light of all circumstances surrounding a data transfer operation or set of data transfer operations.

Awareness Q#9: How will you monitor data transfers?

Data controllers need to identify any transfer of data outside the EU. The handling process must be defined. Data transfer (between whom and whom) within the project, especially with partners from non-EU countries (developed and/or developing countries) must be given special care because of the variety of legal and administrative standards. The EUI's Data Protection Policy as well as EU legislation allow transfers of data outside Europe only to places where there is a local assurance that the level of data protection is equivalent or at least compatible to that of the EU area. Researchers need to consider this aspect not only between institutions and companies, but also within companies and the research partnership across geographical borders.

7 https://www.whitehouse.gov/presidential-actions/executive-order-enhancing-public-safety-interior-united-states/

8. ANONYMISATION

What's in a name?
Anonymisation techniques make data subjects unidentifiable ('further processing')

One of the big advantages of anonymisation is to allow research that would not otherwise be possible due to privacy concerns.

Notions to keep in mind: pseudonymisation vs. anonymisation.

Pseudonymisation still makes the data subject identifiable, through the combination of the pseudonym (e.g. key-code, code number) with additional identifiers. The time and the effort required to identify the individual as well as the available technologies are decisive for determining whether is possible to identify the data subject from pseudonymised data.

Anonymisation excludes any possibility to identify the data subject.

Example: collecting immigrants' data on their immigration experiences could lead to an added value in research on irregular migration, but could also seriously infringe people's privacy and put them at risk of prosecution by the authorities as well as persecution by human smugglers.

A possible solution would be to remove direct identifiers such as names, birth dates, and addresses, although this might not be sufficient to avoid that the data can be traced back to individuals.

An **effective anonymization solution prevents** all parties from singling out an individual in a dataset by:

- linking several records within a dataset (or between several separate datasets)
- inferring any information in such dataset.

Food for thought:

- removing direct identifiers in itself is not enough to ensure that identification of the data subject is no longer possible

- additional measures are usually necessary to prevent identification, depending on the context and purposes of the processing for which the anonymised data are intended.

Example: if a person is described as 'a man' the anonymity set size is three and a half billion, but if he is described as 'a middle-aged Dutchman with a beard' it is maybe half a million and if he is described as 'a middle-aged Dutchman with a beard who lives near Cambridge' it might be three or four.

Data that no longer relate to identifiable persons, such as aggregate or statistical data, are not personal data. Therefore, anonymised data fall outside the scope of data protection rules.

COMMON ANONYMISATION TECHNIQUES

1. Randomisation as…

a family of techniques removing the link between the data and the individual. If the data are sufficiently uncertain then they can no longer relate to a specific individual.

2. Generalisation as…

an approach consisting of generalising or diluting the attributes of data subjects by modifying the respective scale or order of magnitude (i.e. a region rather than a city, a month rather than a week).

Food for thought:
Generalisation can be effective to prevent singling out but does not allow effective anonymisation in all cases. Make sure you devise specific and sophisticated quantitative approaches to prevent linkability and inference.

3. Pseudonymisation as…

a hybrid technique which entails replacing personal identifiers (such as the name) with a unique identifier such as a code.

Main characteristics:

- data subjects are still likely to be indirectly identified but only under pre-defined circumstances
- when used alone it does not result in an anonymous dataset
- it reduces the linkability of a dataset with the original identity of a data subject
- it is a useful security measure but not a method of anonymisation.

Factors influencing its effectiveness:

- stage at which it is used
- its level of protection against reverse tracing
- the size of the population in which the data subject is concealed
- the ability to link individual transactions or records to the same person, etc.

TIPS:

- use random and unpredictable pseudonyms
- make sure the number of pseudonyms possible is so large that the same pseudonym is never randomly selected twice.

Example of pseudonimisation: Key-Coded data.

Common identifiers such as name, date of birth and address could be kept separate from other information on the data subjects, by reference to a code number. Earmarking data subjects by a code would make their identities not immediately apparent. The key that allows the connection between the code and the common identifiers shall be kept separately.

Main characteristics:

- if you use unique codes for each specific person, the risk of identification occurs whenever it is possible to get access to the key used for the encryption. Therefore, codes should be less specific, so that additional information (e.g. year, location for interviews) becomes necessary to 'pierce the veil', i.e. to identify the individual behind the code.

Ultimate Goal: Anonymous data

Anonymous data as…

any information that no longer relates to specific individuals. Data are anonymised as long as the risk of re-identification of the data subjects can be excluded, taking account of all the means likely to be used, either by the data controller or by any other person.

TIPS:

- be cautious as to the timing of the anonymisation process. You are collecting 'anonymised' data only if you implement anonymisation techniques the moment you collect the data for your research. If anonymisation happens at a later stage, (e.g. when you will put the data into a database or during the transcription of audio recordings) you will be processing personal data until the moment of their anonymisation. You will need to provide data protection measures for the raw data

- when you plan to use anonymised datasets, you must specify the source of the datasets (i.e. origins of the data or the manner in which they were obtained).

Food for thought:
Carry out a case-by-case analysis to assess:

- whether the data combined with additional information can allow the identification of an individual
- which means are likely to be used for identification
- whether the data can be considered as anonymous or not

This is particularly relevant in the case of statistical information. The information may be presented as aggregated data, the original sample is not sufficiently large and other pieces of information may enable the identification of individuals.

Awareness Q#10: Have you considered anonymity and confidentiality?

- Clarify whether the data will be anonymised (link to the data subject will be destroyed) or coded (the data will be reversible)
- explain how you will ensure data security and how any link to the research participants will be handled if not fully anonymised
- insist that participants in surveys/experiments use their initials (not ID session numbers or full signatures) when they sign their consent forms, and not simply suggest that they do so
- if the data will not be anonymised, explain why you cannot anonymise the data (e.g. you need to recontact the participants or do follow-up in case of long-term study)
- if the data will be coded, describe the coding system, and who will have access to it, confirm that it cannot be traced back to individuals unless essential for the study
- use data for statistics only after anonymisation techniques have been applied.

9. DELETION AND ARCHIVING OF DATA

When it's over it's over!

You must keep the personal data only as long as it is necessary to fulfil the purpose/s of your research ('retention period'). You can determine specific retention periods also in accordance with the established auditing or archiving for your project. In these cases, you must provide the research participants with explanation of the legitimate basis and purpose for retaining the data, in order to obtain their informed consent.

TIPS:

- securely delete the data once the retention period has expired (i.e. as soon as there are no longer needed for the research purpose/s)
- if you store the data on a cloud system or if a third-party service provider holds the data, make sure that the data will be securely deleted
- if you have transferred the data to third parties, make sure they have deleted them.

LIST OF MAIN REFERENCE DOCUMENTS

- European University Institute – Data Protection at the EUI.
- Horizon 2020 Programme- How to complete your ethics self-assessment (European Commission, Directorate-General for Research & Innovation, Version 6.1, 4 February 2019).
- Ethics and data protection (European Commission, 14 November 2018, here: Ethics and data protection).
- Research and Innovation (European Commission).
- Opinion 05/2014 on Anonymisation Techniques (Art. 29 Working Party).
- Opinion 02/2015 on C-SIG Code of Conduct on Cloud Computing (Art. 29 Working Party).
- Guidelines on consent (Art. 29 Working Party).
- Research Data Management Support Guides (Utrecht University).
- Guidelines on the security and retention of research data (University College Dublin).
- IT Services (University College Dublin).
- Information Commissioner's Office (ICO) website.

ANNEX 1: SAMPLE NOTIFICATION FORM TO BE SUBMITTED TO THE DPO WHEN SEEKING DATA PROTECTION CLEARANCE IN THE CONTEXT OF AN ETHICS REVIEW

EUROPEAN UNIVERSITY INSTITUTE

PROTECTION OF PERSONAL DATA NOTIFICATION OF PROCESSING OPERATIONS - EUI

REFERENCE:
Title: [Please insert the reference number and the title of the project.]

1. PROCESSING	
1.1 Name of the Data Controller	[Please insert the name of the researcher/principal investigator who determines the purposes and means of the processing of personal data.]
1.2 Name of the Processor(s)	[Please indicate the names of any other natural or legal person that may process the data. If processors can be categorised into groups please refer to them by groups and not necessarily by name, otherwise indicate their names.]
1.3 Lawfulness of Processing	[You must process only those personal data that are necessary for your research. Processing personal data that are not essential to your research may even expose you to allegations of 'hidden objectives', i.e. processing information with the data subjects' permission for one purpose and then use that information for another purpose, without specific permission.]
1.4 Description of the processing operations (i.e. what you do with personal data and how)	['Processing of personal data' means any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as: • collection (digital audio recording, digital video caption, etc.) • recording

	• organisation and storage (cloud, LAN or WAN servers) • adaptation or alteration (merging sets, amplification, etc.) • retrieval and consultation • use • disclosure, transmission, dissemination or otherwise making available (share, exchange, transfer, access to the data by a third party) • alignment or combination • blocking, deleting or destruction, etc. Please describe in detail the processing operations that you will perform for conducting your research and give detailed feedback on participants. Indicate also if a copy of notification/authorisation for tracking or observation is required. Any type of research activity may involve processing of personal data (ICT research, genetic sample collection, research activities involving personal records (financial, criminal, education, etc.), lifestyle and health information, family histories, physical characteristics, gender and ethnic background, location tracking and domicile information, etc.)] any method used for tracking or observing.
1.5 Categories of Data Subjects	[Please indicate the categories of data subjects involved in the processing operations of the project.]
1.6 Categories of personal data	[Please list concretely the categories of personal data that you will process: • personal data: name, home address, e-mail address, location data etc. • sensitive data: religious beliefs, political opinions, medical data, sexual identity, etc. • secondary use of data: you must specify if you will process data that were previously collected for another purpose as well as how you will ensure their lawful processing.]

1.7 Rights of data subjects	[Article 16 of the EUI's DP Policy *Data subjects enjoy the following rights concerning their personal data:* *a. to be informed whether, how, by whom and for which purpose they are processed* *b. to ask for their rectification, in case they are inaccurate or incomplete* *c. to demand their erasure in case the processing is un lawful or no longer lawful ('right to be forgotten')* *d. to block their further processing whilst the conditions under letters b) and c) of this Article are verified.* Please indicate how you will ensure the data subjects' rights. E.g. participants will be free to withdraw at any time without justification. The data collected prior to the withdrawal will be deleted. In such a case, you may need to ensure the erasure of the collected data while maintaining anonymity. In order to do so, you may use a pseudonym for each participant ensuring that the key to the pseudonyms is password-protected and available only to the data controller.]
2. DETAILED PROCEDURES	
2.1 Details on the procedures that you will use to identify/ recruit research participants.	[A few examples: • consultative process involving a gatekeeper (e.g. NGOs providing support to the participants) • snowballing (through referral from one participant to another) • personal contacts (obtained by the researcher from the contextual knowledge of the country and place where the research is conducted) • etc.]
2.2 Details on the procedures for obtaining informed consent	[Please give details on the procedures for obtaining informed consent from the data subjects (e.g. providing an information sheet together with the consent form). In case of children/minors and/or adults unable

	to give informed consent, indicate the tailored methods used to obtain consent. According to the H2020 Guidelines, if the data subjects are unable to give consent in writing, for example because of illiteracy, the non-written consent must be formally documented and independently witnessed. Please explain how you intend to document oral consent. In the very exceptional case that it can't be recorded please give reasons. If you will use deception for another type of data subjects, you must obtain retrospective informed and free consent as well as debrief the participants. Deception requires strong justification and appropriate assessment of the impact and the risk incurred by both researchers and participants.]
2.3 Measures taken to prevent the risk of enhancing vulnerability/ stigmatisation of individuals/groups	[Please indicate any such protective measures (e.g. use of anonymisation techniques, use of pseudonyms, non-disclosure of audio-visual materials, voice records, etc.)]
2.4 Safeguards taken to protect the data subjects' identity.	[Article 2 of the EUI's DP Policy Identifiable persons can be identified directly or indirectly, in particular by reference to an identification number or to one or more factors specific to their physical, physiological, genetic, mental, economic, cultural or social identity. Please provide details on the measures taken to avoid direct or indirect identification of the data subjects, e.g. by using anonymisation techniques or pseudonyms. E.g. names of the data subjects will not be disclosed, at any time, in audio recording and published material. Pseudonyms (a reversible system of coding in order to be able to re-contact participants if needed) will be used in all documentation, and any additional information that may reveal the

	identity of participants will be concealed when publishing. Destroy any residual information that could lead to the identification of participants at the end of the project. You must explain this procedure clearly to participants during the 'recruitment' process].
3. DETAILED PROCEDURES You need to work out and clearly define a secure access policy. It must be proportional to the risks involved and to the sensitivity of the data. Please state clearly the type of security measures – such as password protection, encryption, 'need to know basis' principles (i.e. only the users that need to access the data will be allowed to do so), - that you will implement.	
3.1 Storage medium	[Please indicate any methods considered for data storage. E.g.: You will identify each transcription by a pseudonym whilst you will store the data subjects' names in a separate file to ensure security. You will password-protect all of these files. You will store data collected in a secure database at the EUI, to which only the project team members will have access, in order to prevent any unauthorised access and possible misuse (e.g. data mining, profiling). E.g.: Temporary storage (on site): The transcribed interviews and field observations will be stored electronically and password-protected. The researcher will make regular back-up copies of these files, which will be stored offline on the hard
3.2 Retention Period	[Article 4 of EUI's DP Policy *Personal data are kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which they were collected or for which they are further processed.* Please indicate how you will comply with this requirement by establishing the exact retention period as well as the measures that you will implement to delete the data.]

4. DATA TRANSFER	
4.1 Within the EUI	Will you transfer any personal data **to recipients inside the EUI**? To whom (person(s) or category of person(s)/legal entities) & for what purpose will you transfer the data?
4.2 Inside the EU/EEA	Will you transfer any personal data **to third parties inside the EU/EEA area**? To whom (person(s) or category of person(s)/legal entities) & for what purpose will you transfer the data?
4.3 Outside the EU/EEA	Will you transfer any personal data to **third parties outside the EU/EEA area**? How will you ensure that the third party has in place adequate safeguards for the protection of personal data (e.g. adequacy decision adopted by the European Commission?)?
5. COMPLEMENTARY INFORMATION If necessary	

ANNEX II: SAMPLE CONSENT FOR PARTICIPATION IN RESEARCH INTERVIEW

Consent for participation in research interview

[name of the project]

funded by

[name of the sponsor]

I agree to participate in a research project conducted by Prof. *[Name of the Principal Investigator]* from the European University Institute (EUI) in Florence, Italy.

1. I have received sufficient information about this research project and understand my role in it. The purpose of my participation as an interviewee in this project and the future processing of my personal data has been explained to me and are is clear.
2. My participation as an interviewee in this project is completely voluntary. There is no explicit or implicit coercion whatsoever to participate.
3. Participation involves being interviewed by (a) researcher(s) from the European University Institute. The interview will last approximately *[xxx]* minutes. I allow the researcher(s) to take notes during the interview. I also may allow the recording of the interview and subsequent dialogue by audio/video tape. It is clear to me that in case I do not want the interview and dialogue to be taped I am fully entitled to withdraw from participation.
4. I have the right not to answer questions. If I feel uncomfortable in any way during the interview session, I have the right to withdraw from the interview and ask that the data collected prior to the withdrawal will be deleted.
5. I have been given the explicit guarantee that the researcher will not identify me by name or function in any reports using information obtained from this interview, that my confidentiality as a participant in this study remains secure. Personal data will be processed in full compliance with the EUI's Data Protection Policy.
6. I was assured that this research project has been reviewed and approved by *[xxx and]* by the EUI Ethics Committee. The EUI Ethics Committee may be contacted through *[information of the contact person at the Ethics Committee at EUI]* for any questions concerning ethics. The Data Controller may be contacted at *[e-mail address of the Data Controller]* for any questions concerning data protection. Complaints should be addressed to the Data Controller with copy to the EUI's Data Protection Officer (data_protection_officer@eui.eu).
7. I have carefully read and fully understood the points and statements of this form. All my questions were answered to my satisfaction, and I voluntarily agree to participate in this study.
8. I obtained a copy of this consent form co-signed by the interviewer.

_____ _____

Participant's Signature Date

_____ _____

Researcher's Signature Date

For further information, please contact:
Prof. *[Name of Principle Investigator* – Data Controller*]* at *[contact information of PI]*

Version: June 2022