EJP RD
European Joint Programme on Rare Diseases

H2020-SC1-2018-Single-Stage-RTD
SC1-BHC-04-2018
Rare Disease European Joint Programme Cofund

Grant agreement number 825575

GDPR Guiding Technical Principles for EJP RD Pillar 2

Disclaimer: *This document does not replace legal advice, nor it is meant to be covering all possible situations and options available. It is meant as a simple introductory guide for all contributors to the EJP RD Pillar 2 to facilitate orientation in the data protection domain.*

Authors: *Petr Holub, Irene Schlünder*

Table of contents

1. Possible Data Protection Roles

When processing data subject to data protection regulations (further denoted just as personal data), you are likely to assume one of the following roles - ordered roughly from the least sensitive with respect to the data protection obligations:

- **I am a software provider.** You don't have any particular role with respect to GDPR. However, if your software deals with personal data, you want to support adopters of your software by *adopting "privacy by design"*[1] in your software. In particular if your service is likely to work with data of special categories[2], adopters of your software are likely to perform Data Protection Impact Assessment (DPIA)[3] and you can facilitate this process by providing documentation on how technical measures for data protection are implemented in your software.
- **I am a service provider, but my service gets deployed by data controllers/processors.** Dtto as a software provider.
- **I am an infrastructure provider where my users could deposit strongly encrypted data and I don't have keys to decrypt those.** You may consider the data being just "binary mess" and if the user is OK with depositing the data at your infrastructure, you can consider those non-sensitive.
- **I am a service provider and have my users uploading their data to my service to have them processed by my service.** You have a role of processor and you should have a contract with the users uploading the data that allows data processing and determines the security measures.
- **I am a maintainer of a database containing personal data.** Depending on the situation you can be *a data controller or a processor*. Whereas the controller decides on for what purpose and how the data are processed (see below), the processor follows the instructions of the controller. In many situations both options are possible technically and you need to *decide* which is best considering where the data comes from and what is the anticipated life cycle

1 GDRP Art. 25 and GDPR Rec. 78
2 GDPR Art. 9
3 GDPR Art. 35

of the data. For the processor role it is often [implicitly] assumed that processing is time limited, i.e., the controller can decide to transfer the data to another processor. Please note that in particular Finland does not currently allow transfer of certain data to another controller outside of Finland - which may limit your options if Finnish data is to be included[4]. If you become a processor, your duties and restrictions on data processing must be specified in the contract you receive from the data controller or upstream data processor.

- **I am a maintainer of a database containing personal data of special categories[5].** Dtto as above, plus you are typically required to perform Data Protection Impact Assessment (DPIA)[6].

2. General Notes

- Establishing a new data controller[7] based on controller-to-controller contract does not presume that the original controller would cease to be a controller; i.e., both the original controller and the new controller become controllers in such a case usually. Establishing a new controller may result in two independent controllers or in joint controllership (which is least advisable as discussed above).
- If you need/want to sign any data protection related contracts, make sure you know *who is the person legally entitled to sign the contract* (in many cases it is not the researcher who wants to have the contract in place, but rather a legal representative of the institution). This applies for both (all) parties signing the contract.
- *Processing of personal data should be registered at the institution*, so that each institution has a reasonable overview where data protection issues are relevant.
- Make sure you *consult the Data Protection Officer* at your institution.
- The responsibility related to data controllership is institutional responsibility, not the responsibility of individual person. Therefore, it is important to understand that the individuals are working on behalf of their home institution, based on delegation established as a part of their employment.

3. Notes of Data Controllership

- **Who is controller?** It is the entity that decides on the purpose of the processing and defines means of processing[8]. Note it does not matter at the court what is

4 This is witnessed by (central) EGA, dbSNP, as well as BBMRI-ERIC CRC-Cohort.
5 GDPR Art. 9
6 GDPR Art. 35
7 Establishing a new controller is sometimes denoted as "transfer of data" based on the fact that it is usually governed by "data transfer agreement", which deals with all the different aspects of how the data is transferred from one entity to the other, including specifying role of the recipient of the data.
8 GDPR Art. 4

written on the paper, but what is reality - so as long as you act as a controller, you are a controller with all the duties related to controllership.

- **Transfer of data.** The mere fact that you receive data from somebody does not directly imply that you are a data processor on their behalf. It also does not imply that you become joint controllers (see below). The data transfer means you can become a data controller or a data processor, depending on your needs, the purpose of the transferring party and thus on the type of contract you have between you and the upstream data provider.
- **Being a controller while outsourcing data collection process onto a processor.** This may seem a bit counterintuitive, but a data controller can use a service of a third party to collect the data, while still having this third party only a data processor on behalf of the controller. This can be used in situations where the controller does not have personal capacity to collect the data.
- **Joint controllership.** This is rather problematic situation. In practice, joint controllership is never recommended in legal books, since joint data controllers bear jointly all the risks and liability. Hence it is the least advisable setup of controllership; if the goals can be achieved by transfer of data from controller to controller (thus having a sequence of controllers) or by controller-processor setup, it is always preferred.

4. Notes on Anonymization

Anonymized data means data to which anonymization has been applied and residual risk left after anonymization has been assessed. Anonymization means a complex process which manipulates the data in such a way that data subjects are not distinguishable within the anonymity set and also that risks of other privacy attacks is minimized (e.g., risk of group membership inference or inference of additional data based on external knowledge of group membership). Anonymization is not a binary process and there is no "perfectly anonymized data", which is still useful for research purposes (c.f. Section 1.1 of The Algorithmic Foundations of Differential Privacy[9]); there is always some residual risk left and this risk needs to be (re)assessed periodically if acceptable.

5. Process Workflows

Here are the process workflows describing the most common situations:

- starting a new research project;
- implementing data management phase during the project and after the project is over;

9 Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy." *Foundations and Trends® in Theoretical Computer Science* 9.3–4 (2014): 211-407. https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf

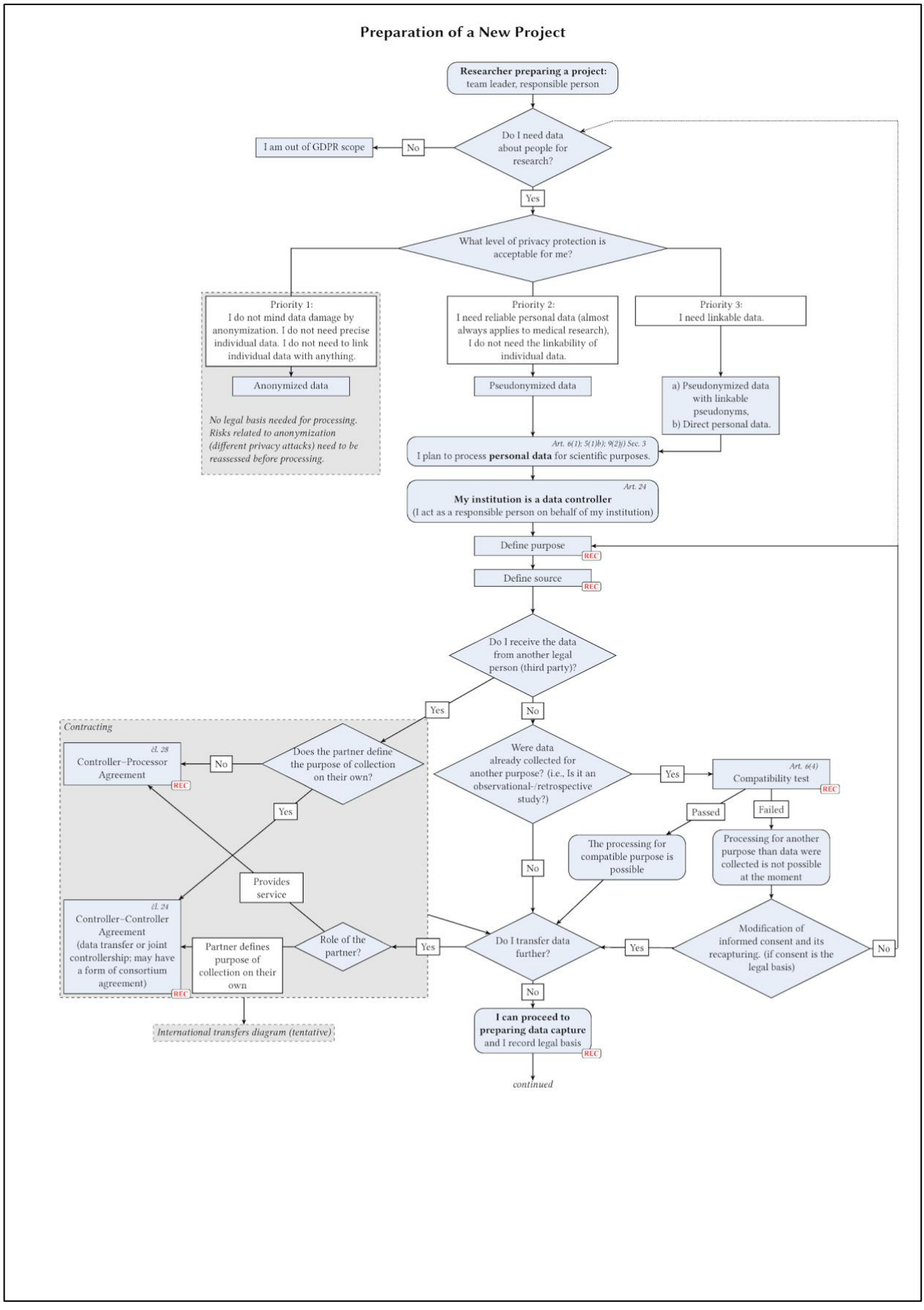
- implementing reactions to events and how to utilize exemptions provided by GDPR for research data.

These diagrams adapted from the Czech recommendation how to deal with research data under GDPR[10]. The have been updated based on current state of the art.

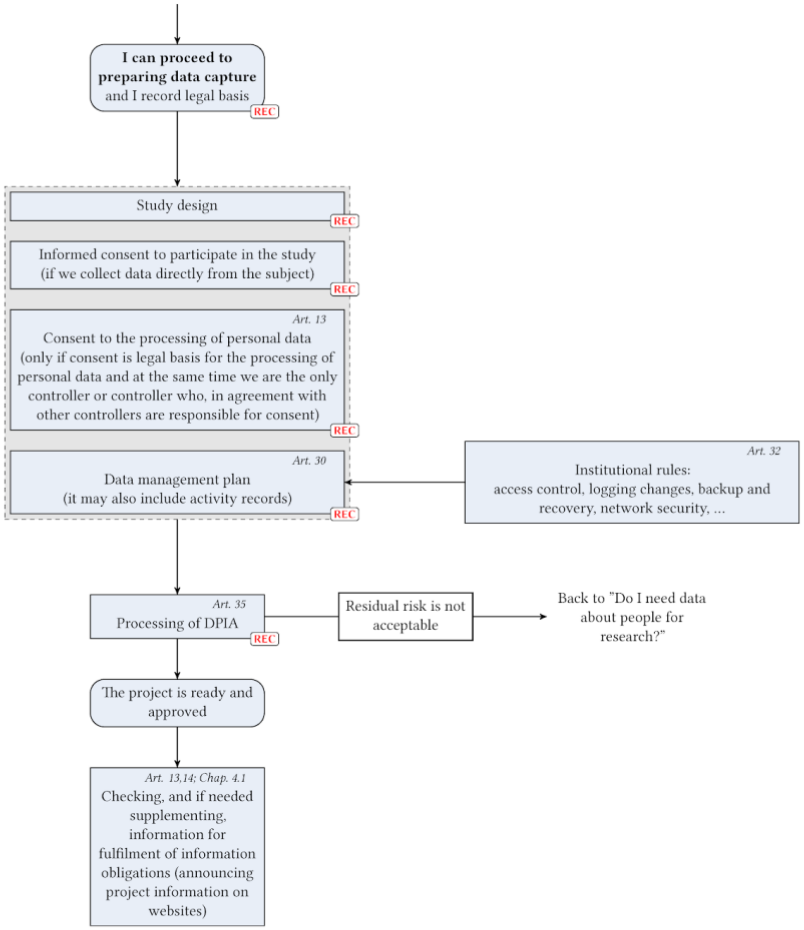
Note that the diagram on setting up a new project uses rough approximation that anonymized data is outside GDPR as it is no longer personal data; however, because of the residual risk it is not completely true and anonymization of the data needs to be (re)assessed as explained above.

The part of the diagram describing international transfers is very tentative because of rapidly developing situation on the international scene.

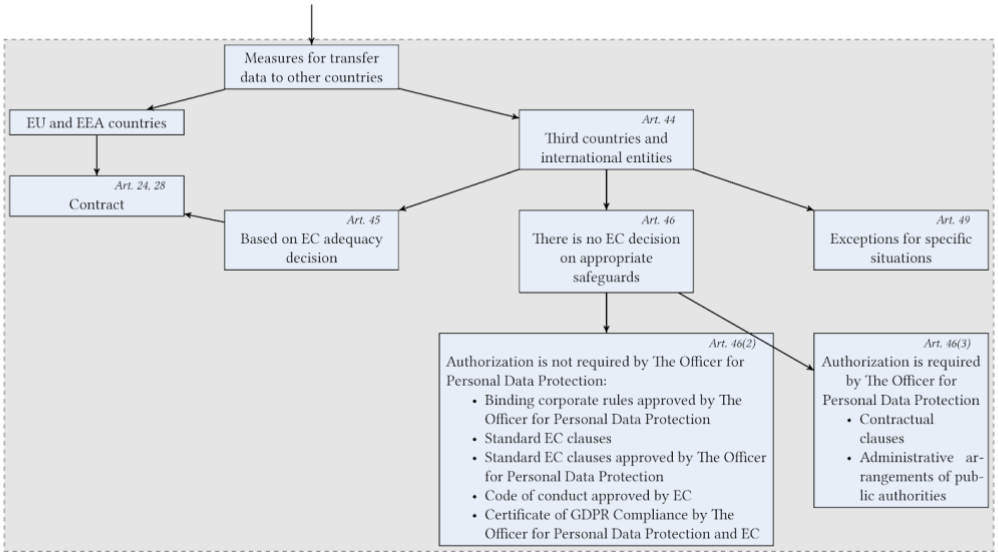
[10] Radim Polčák, Leoš Ševčík, Michal Koščík, Jakub Klodwig, & Petr Holub. (2018). Metodika aplikace GDPR na výzkumná data v prostředí vysokých škol v ČR. Zenodo. https://zenodo.org/record/2533865

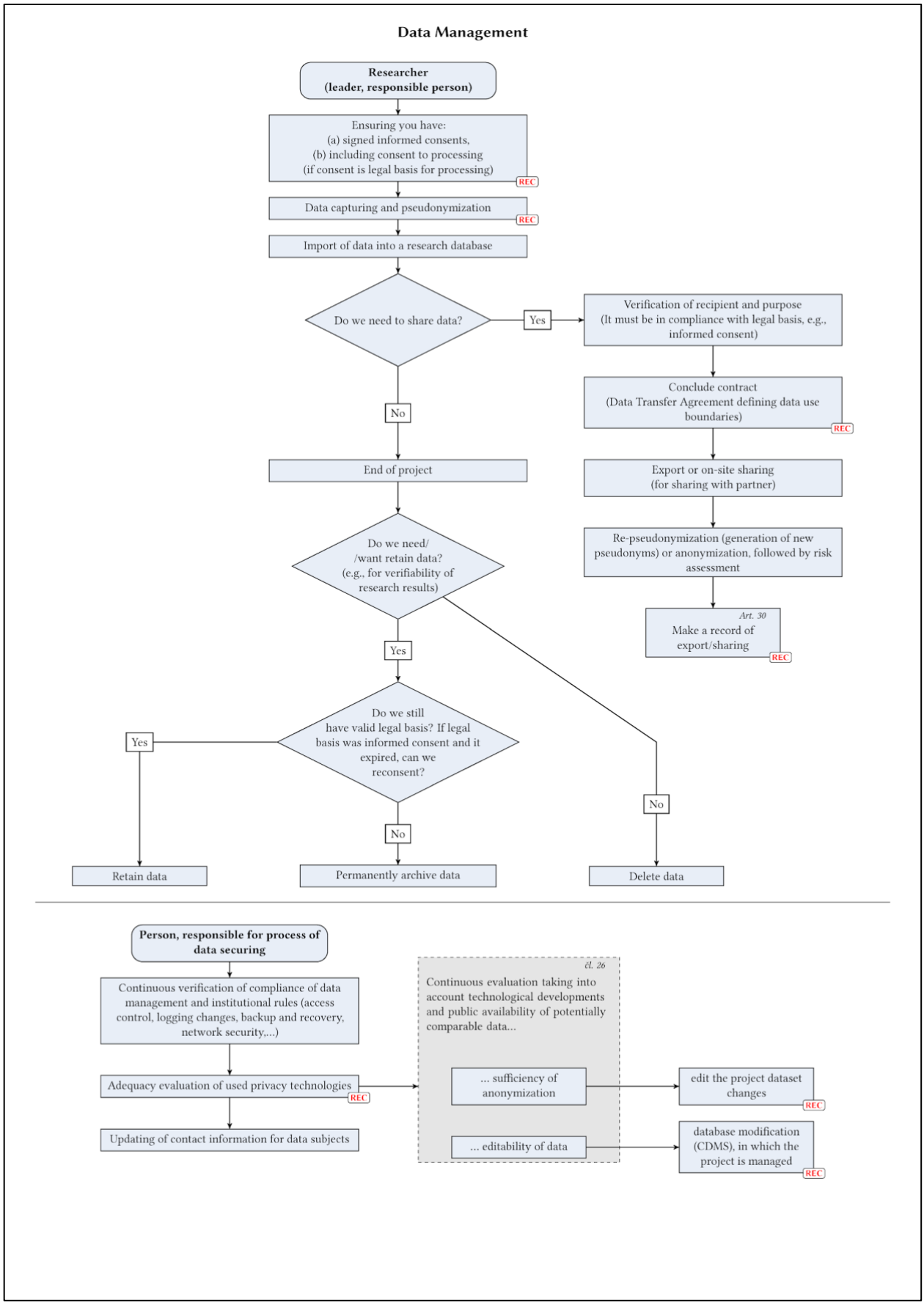

Preparation of a New Project
Researcher preparing a project: team leader, responsible person
Do I need data about people for research?
No
I am out of GDPR scope
Yes
What level of privacy protection is acceptable for me?
Priority 1: I do not mind data damage by anonymization. I do not need precise individual data. I do not need to link individual data with anything.
Anonymized data
No legal basis needed for processing. Risks related to anonymization (different privacy attacks) need to be reassessed before processing.
Priority 2: I need reliable personal data (almost always applies to medical research), I do not need the linkability of individual data.
Pseudonymized data
Priority 3: I need linkable data.
a) Pseudonymized data with linkable pseudonyms, b) Direct personal data.
Art. 6(1); 5(1)b); 9(2)j) Sec. 3
I plan to process personal data for scientific purposes.
Art. 24
My institution is a data controller (I act as a responsible person on behalf of my institution)
Define purpose
REC
Define source
REC
Do I receive the data from another legal person (third party)?
Yes
No
Contracting
čl. 28
Controller–Processor Agreement
REC
No
Does the partner define the purpose of collection on their own?
Yes
Were data already collected for another purpose? (i.e., Is it an observational-/retrospective study?)
Yes
Art. 6(4)
Compatibility test
REC
Passed
Failed
The processing for compatible purpose is possible
Processing for another purpose than data were collected is not possible at the moment
No
Provides service
čl. 24
Controller–Controller Agreement (data transfer or joint controllership; may have a form of consortium agreement)
REC
Partner defines purpose of collection on their own
Role of the partner?
Yes
Do I transfer data further?
Yes
Modification of informed consent and its recapturing. (if consent is the legal basis)
No
No
I can proceed to preparing data capture and I record legal basis
REC
International transfers diagram (tentative)
continued

Preparation of a new project (continued)



International transfers diagram (tentative)




Data Management
Researcher
(leader, responsible person)
Ensuring you have:
(a) signed informed consents,
(b) including consent to processing
(if consent is legal basis for processing)
REC
Data capturing and pseudonymization
REC
Import of data into a research database
Do we need to share data?
Yes
Verification of recipient and purpose
(It must be in compliance with legal basis, e.g., informed consent)
Conclude contract
(Data Transfer Agreement defining data use boundaries)
REC
Export or on-site sharing
(for sharing with partner)
Re-pseudonymization (generation of new pseudonyms) or anonymization, followed by risk assessment
Art. 30
Make a record of export/sharing
REC
No
End of project
Do we need/
/want retain data?
(e.g., for verifiability of research results)
Yes
Do we still
have valid legal basis? If legal basis was informed consent and it expired, can we reconsent?
Yes
No
No
Retain data
Permanently archive data
Delete data
Person, responsible for process of data securing
Continuous verification of compliance of data management and institutional rules (access control, logging changes, backup and recovery, network security,...)
Adequacy evaluation of used privacy technologies
REC
Updating of contact information for data subjects
čl. 26
Continuous evaluation taking into account technological developments and public availability of potentially comparable data...
... sufficiency of anonymization
... editability of data
edit the project dataset changes
REC
database modification (CDMS), in which the project is managed
REC