EPFL

LEGAL AFFAIRS

Data Protection in Research Projects Guidelines

École
polytechnique
fédérale
de Lausanne

EPFL P-SG AJ
CE 3 319 (CENTRE EST)
STATION 1
CH-1015 LAUSANNE

+41 21 693 45 31
chiara.tanteri@epfl.ch
www.epfl.ch

Version - November 2021

Table of contents

#1 Purpose of this document

The purpose of these Guidelines is to provide the EPFL research community a resource to help it to manage the personal data correctly involved in their research projects, to protect the privacy of the data subjects and to ensure high data protection and scientific standards.

This document aims to present, at each stage of the data life cycle, not only concepts or legal reminders but concrete examples from research and who or which EPFL entity to contact in case of need.

We are convinced that the protection of personal data is not a constraint to be respected but a means and an opportunity to strengthen citizens' confidence in science

1.1. Why is the law on data protection so important?

The protection of an individual's fundamental rights and freedoms is laid down in the Constitution. In this sense, the title of the Federal Act on Data Protection is simplistic.

This law effectively aims to protect the fundamental rights and privacy of individuals, by regulating the processing of their personal data.

Any processing of data which does not respect the principles laid down by the law on data protection can cause physical, material or immaterial damages, or discrimination (e.g. disclosure of data, reception of spam, unanticipated payments, abusive profiling, defamation or reputational damages, intimidation on social networks, financial loss following fraud, psychological harassment, discrimination, etc.).

In a world where online services are highly developed, and where the quantity of personal data in circulation on the Internet is increasing at incredible speed, it is essential for individuals to be able to trust the public and private actors who process their data.

The law on data protection must not be considered as an obstacle to development; the creation of a climate of trust within the online environment is essential to this development, whether scientific or economic.

As a federal body, EPFL must respect the Constitution and applicable laws.

1.2. Authors

This document has been drafted in 2021 by a dedicated working group, coordinated by the Data Protection Officer.

Working group members:

- G. Anex (SV-IT)
- M. Braskova (STI-IT)
- E. Blumer (VPA-SISB-GE)
- V. Conrad (VPA-EM-AJ)
- G. Dubochet (VPA-OS-GE)
- J.-F. Dousson (GOUV-GE)
- Z. Khaliliardali (VPA-EM-ACAD)
- S. Kury (CIGR)
- C. Tanteri (AJ)
- E. van der Velde (AVP-R-REO)
- M. Wessel ()

Coordination: C. Tanteri

Graphic layout: Mediacom Communication Visuelle (MCV)

#2 The main definitions

In this chapter, the main data protection definitions are given as well as some examples.

The Federal Act on Data Protection[1] defines the most important data protection concepts.

2.1. Personal data

Personal data is all information relating to an identified or identifiable person.

It is not so important what form personal data takes - it may be a sign, a writing, an image, a sound or a combination of these elements[2].

→ *Advanced – Main definitions - part 1*

Examples of directly identifying personal data:

- name
- address
- photo
- voice

Examples of indirectly identifying personal data:

- telephone number
- SCIPER number
- AVS number
- location data
- Internet Protocol (IP) address

2.2. Data subjects

These are the natural or legal persons whose data is processed.

Examples of data subjects:

- the human participants that take part in the research project.

There are different categories of human participants, such as:

- patients
- healthy volunteers
- children or adolescents
- vulnerable adults (such as pregnant women, prisoners)
- populations in developing countries

2.3. Sensitive personal data

This is a sub-category of personal data and it includes data on: (1) religious, ideological, political or trade union-related views or activities; (2) health, the intimate sphere or the racial origin; (3) genetic data; (4) biometric data which unequivocally identifies a natural person; (5) data on administrative or criminal proceedings and sanctions; (6) social security measures.

Examples of sensitive personal data:

- biometric data
- medical images
- hospital records
- biological traits and genetic information
- membership of a political party or religious group
- sexual orientation
- criminal records

[1] The Federal Act on Data Protection has been revised on 25 September 2020. The revised version, that will enter into force by the end of 2022, has been used for this guidelines.

[2] JdT 2010 I 396

2.4. Profiling

The profiling encompasses the entire process that leads to the creation of a personality profile[3]. It includes any form of automated processing of personal data consisting of using such data to assess certain personal aspects relating to a natural person, in particular to analyse or predict aspects relating to the performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or whereabouts.

A profiling is considered as a high risk profiling, when it involves a high risk to the personality or fundamental rights of the data subject, as it creates a pairing between data that enables an assessment of essential aspects of the personality of a natural person.

2.5. Processing

Processing covers a wide range of operations performed on personal data. In legal terms it means "any operation with personal data, irrespective of the means applied and the procedure, such as the collection, storage, use, revision, disclosure, archiving or destruction of data".

Examples of processing:

the automatic collection of personal data on the web, the management of a database, sharing data with a third party, video recording of human participants in a research project, storing IP addresses or MAC addresses, the creation of a mailing list of participants, the anonymization of data, etc.

2.6. Data controller vs data processor

The FADP defines the controller of the data file as follows: "private person or federal body that alone or jointly with others decides on the purpose and the means of the processing".

The data controller determines the purpose for which and the means by which personal data is processed. An institution/organization can also be a joint controller when together with one or more organizations it jointly determines why and how the personal data should be processed.

The 'purpose' relates to the goal of the processing: why is the data collection is needed? The 'means' refers to the essential characteristics of the processing design: what kind of personal data needs to be collected? For how long will the data be stored? To whom will the data be communicated?

A data processor processes the data on behalf of the controller. The data processor is an external entity, mostly a service provider. The data processor is not able to change the purpose and the means of the use of the data, it is bound by the instructions it received from the data controller. It cannot use the data for its own purposes. For example, a research lab is tasked to perform an analysis on data that is provided by another institution (the data controller). It will then return the data (and the results) to that institution.

The duties of the processor towards the controller must be specified in a written contract or agreement[3].

[2] Cellina Eva, La commercialisation des données personnelles, Schulthess, 2020, § 174.

[3] https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor_en

2.7. Anonymization vs Pseudonymization

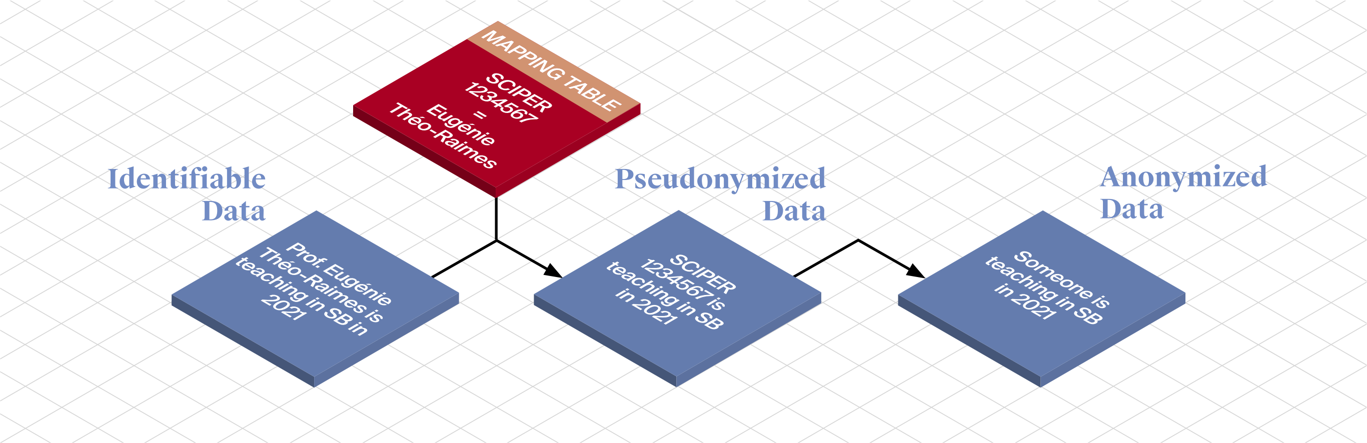
According to the Federal Data Protection Commissioner, personal data is pseudonymized when it is replaced by a code (pseudonym), while it is anonymized when all identifying data is removed. Pseudonymization is reversible while anonymization is definitive[4].

Irreversibly anonymized data, which no longer allow the re-identification of a person, are not subject to the regulations on the protection of personal data[5].

Note: The Human Research Act (HRA) provides definitions for coded and uncoded health-related personal data in Art. 3. Following the guidance of the Swiss Academy of Medical Sciences following degrees of anonymization can be distinguished.

For research projects for which swissethics is the responsible authority, accepted coding formats have been formalized (link: https://www.swissethics.ch/en/themen/weitere-themen). Coding numbers containing year of birth are accepted. On the contrary, coding containing data of birth and/or initials of the participants are not accepted.

→ ***Advanced – Main definitions - part 2***


MAPPING TABLE
SCIPER 1234567 = Eugénie Théo-Raimes
Identifiable Data
Prof. Eugénie Théo-Raimes is teaching in SB in 2021
Pseudonymized Data
SCIPER 1234567 is teaching in SB in 2021
Anonymized Data
Someone is teaching in SB in 2021

2.8. Open Research Data

Research data (RD), including code, is defined as evidence that underpins the answer to the research question, and can be used to validate findings regardless of its form (e.g. print, digital, or physical).

Open research data are one part of RD that can be freely accessed, used, modified, and shared, provided that there is appropriate acknowledgement if required[6].

As a matter of fact, not all research data can be open and it is commonly recognized that access may need to be managed to maintain confidentiality, guard against unreasonable cost, protect individuals' privacy, respect consent terms, raise possible intellectual property issues, as well as manage security or other risks.

[4] Préposé fédéral à la protection des données et à la transparence PFPDT, Guide relatif aux mesures techniques et organisationnelles de la protection des données, août 2015, p. 15

[5] The processing of anonymizing the data is however subject to the regulations on data protection.

[6] Concordat on Open Research Data: https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-

#3 The main stakeholders and EPFL support units

This section outlines the main stakeholders in the context of personal data.

In a next step, this section highlights the different support units at EPFL when it comes to handling personal data in research. It explains the role of the different units and what support they can provide for research projects involving personal data.

3.1. Main stakeholders

Principal Investigator (ICH E6, 1.34)

If a trial is conducted by a team of individuals at the trial site, the investigator who is the responsible leader of the team may be called the principal investigator.

Sponsor (ICH E6, 1.53)

An individual, company, institution, or organisation which takes responsibility for the initiation, management, and/or financing of a clinical trial.

Sponsor-Investigator (ICH E6, 1.54)

An individual who both initiates and conducts, alone or with others, a clinical trial, and under whose immediate direction the investigational product is administered to, dispensed to, or used by a subject. The term does not include any person other than an individual (e.g., it does not include a corporation or an agency). The obligations of a sponsor/investigator include both those of a sponsor and those of an investigator.

Researchers

An individual, company, institution or organization which is the operational lead of a study with human subjects and reports to the principal investigator.

Human Subject / Research Participant

According to 45 CFR 46, a human subject is «a living individual about whom an investigator (whether professional or student) conducting research.

Reference:
https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice

3.2. EPFL Support units and their missions regarding personal data

VPA Legal Affairs

Mission:

- Data sharing and processing agreements for researchers who receive or collect data
- Review of research projects that are submitted to the Human Research Ethics Committee (HREC) (3 Legal Counsels are members of the HREC)
- Advise on data protection matters in relation with a research project

Contact: research@epfl.ch

Research Office

Mission:

- Informs the EPFL research community on funding opportunities and assists researchers in drawing up funding applications and research agreements.
- Provides support to researchers with the administrative management throughout their research projects.
- With regard to personal data, the ReO-Ethics Affairs provides support to researchers that have ethical issues in their projects (in particular with regard to ethical approval processes), conducts ethics compliance reviews of research projects and is the liaison with the EPFL Human Research Ethics Committee (HREC).

Contact: research@epfl.ch

Library

Mission:

- Prepares, reviews and evaluates Data Management Plans
- Proposes appropriate data licenses
- Supports and educates on general data privacy good practices
- Offers an archiving tool and support service
- Website with extensive information
- RDM Team training offer around research data management (on demand and scheduled)
- DMP Planning and Systematic Reviewing Grid for any kind of data management plan in collaboration with the existing central and faculty-services
- General support on all questions of research data management (such as metadata, documentation, ELN tools, publication and archiving) and coordination between the existing EPFL services
- Hosting and evolution of the Data Champions Community
- Management of the EPFL Zenodo Community for data publication
- Curated and commented list of tools existing at EPFL for each step of the data life cycle (in elaboration : data repository tools used at EPFL)
- Data Archiving Service ACOUA

Contact: researchdata@epfl.ch

VPO-SI

Mission:

- Sets up and manages data security systems
- Implements digitization projects
- Plans, develops and maintains the School's IT assets

Contact: 1234@epfl.ch

Faculty-ITs:

Mission:

- Provide services to meet the IT requirements of their faculty and for faculty-specific technology, i.e. RedCap

Contacts:

- ENAC-IT: enacit@epfl.ch
- SV-IT:
  - *helpdesk.sv@epfl.ch*
  - *it.ge@epfl.ch*
  - *it.vs@epfl.ch*
- STI-IT: help-sti@groupes.epfl.ch
- SB-IT:
  - *it.iphys@epfl.ch*
  - *it.isic@epfl.ch*
  - *it.math@epfl.ch*
  - *spc.it@epfl.ch*
- IC-IT:
- CDM-IT: support-cdmit@epfl.ch

Data Protection Officer (DPO)

Mission:

- Accompanies EPFL in its compliance with laws and regulations on personal data and with the mitigation of risks related to non-compliance that can have important consequences (finances/subsidies, fines, reputation,...)
- Advises in case of Data Protection Impact Assessment (DPIA) & monitor its performance
- Acts as contact point for requests from individuals about the processing of their personal data and the exercise of their rights
- Cooperates with DP Authorities (DPAs) and acts as a contact point for DPAs on issues on data processing
- Presents an annual report directly to EPFL Management

Contact: dpo@epfl.ch

Open Science Advisor

Mission:

- Supports EPFL President and Provost and Open Science Strategic Committee in definition of EPFL Open Science strategy
- Represents EPFL in national and international Open Science bodies ETH Domain ORD task force, Delegation Open Science (DelOS)
- Recommends and coordinates the development of an Open Science Service ecosystem across EPFL
- Gives advice on the implementation of services / policies w.r.t. Open Science DMPs, Good research practices, School Open Science policies
- Liaises with researchers involved in Open Science practices Management of Open Science Fund
- Proposes communication, information, training, events (Open Science website)

Contact: gilles.dubochet@epfl.ch

In case a private company is involved in your research project, or if you deal with intellectual property rights, please contact the **TTO Office**: info.tto@epfl.ch

#4 The applicable legal framework

In this chapter, the legal framework that applies to research involving personal data conducted by EPFL researchers is explained.

4.1. Swiss Federal Data Protection Act (FADP)

The main data protection law applicable to EPFL is the Federal Act on Data Protection (FADP), supplemented by an Ordinance to the Federal Act on Data Protection (OFADP), which makes certain aspects of the FADP more concrete. The FADP of 1992 has been revised in 2020 and the revised FADP shall enter into force by the end of 2022. For the sake of clarity, this guide is already based on the revised law.

The purpose of the FADP is to protect the privacy, interests and fundamental rights of data subjects and it regulates the processing of personal data. Sensitive personal data (such as data relating to religion, political beliefs, trade union activities, health, etc.) are subject to a higher level of protection (see further down below for more details). The FADP does not apply to the processing of anonymous data (i.e. information or data cannot be linked to the person anymore.)

The Swiss Federal Data Protection and Information Commissioner (FDPIC) supervises the compliance of Federal authorities with the FADP.

In addition to the FADP, the cantons in Switzerland have their own data protection legislation that apply to the cantonal bodys.

Sectorial laws may also supplement and derogate to the FADP. In the field of research, the HRA applies when researchers process health data in connection with a study or clinical trial (for instance).

What will the main changes be* ?

Current FADP	GDPR	Revised FADP
Protection of privacy for natural and legal persons	Protection of privacy for natural persons	Protection of privacy for natural persons only (Art. 1)
Teritorial and limited extraterritorial application	Territorial and extraterritorial application (Art. 3, Establishment, intention/singling out, monitoring)	Territorial and **extraterritorial** application. Extension of the field of application to cover actions which take place abroad and which produce effects in Switzerland (Article 3). Obligation to appoint a representative under certain conditions (Art. 14 and 15).
Sensitive data	Sensitive data, including biometric and genetic data	Sensitive data, including **biometric** and **genetic data** (Art. 5)
Duty to declare data files in certain situations	Records of processing activities (mandatory from 250 employees or if the processing presents a risk)	Records of processing activities (mandatory from 250 employees or if the processing presents a risk), with exceptions for the Federal Council (Art. 12)
	Privacy by default and by design	**Privacy by default and by design**
	Notification in case of breach	**Notification** in case of breach
	Impact analysis (self-regulation, mandatory in certain cases)	**Impact analysis** (mandatory in certain cases)
Duty to provide information on the collection of sensible personal data and personality profiles	Duty to provide information on the collection of all data	Duty to provide information on the collection of all data, duty to provide information on the name of the State in situations of communication abroad (Art. 19 and 20)
Right of access an obligation to provide information on the collection of all data for federal entities and a right of deletion.	Right of access , of deletion	Right of access (**strengthened and specified**) and to request deletion
	Right to data portability	Right to receive or transfer personal data (data **portability**) (subject to acceptions) (Art. 28 and 29)
	Right to not be subject to automated decision-making and duty to provide information	Right to not be subject to **automated decision-making** and duty to provide information

What are the stakes and sanctions* ?

Current FADP	GDPA	Revised FADP
Criminal sanctions: max. **CHF 10,000.-**	Administrative sanctions (max. **EUR 20 million or 4% of the annual global turnover**) The amount to be paid depends on various criteria	Criminal sanctions (private parties => natural) (by the cantons). Fine of **CHF 250,000.– max.** (willfully, or recklessly) › In case of inaccurate or incomplete details relating to the duty to provide information and the right to access or failure to provide information (Art. 60). › In case of communication abroad without the conditions having been met › In case of subcontracting without meeting the conditions › In case of non-respect of the minimum security requirements (Art. 61) › n case of non-conformance to a decision of the Federal Data Protection and Information Commissioner (Art 63) Criminal sanctions (anyone) (by the cantons) Breach of the duty of secrecy CHF 50,000 max for companies (instead of natural persons) (Art. 64) Denunciation by the Federal Data Protection and Information Commissioner (Art 63)
Legal action - > disruption of project, DI	**Legal** action - > disruption of project, DI	**Legal** action - > disruption of project, DI
Reputation of the company, bad publicity (publication of the ruling), loss in value for the company, advantageous for **competitors**	**Reputation** of the company, bad **publicity** (publication of the ruling), loss in value for the company, advantageous for **competitors**	**Reputation** of the company, bad publicity (publication of the ruling), loss in value for the company, advantageous for **competitors**
Loss of clients, partnerships, investors, etc.	**Loss** of clients, partnerships, investors, etc.	**Loss** of clients, partnerships, investors, etc.

4.2. Federal Act on Research involving Human Beings (or: Human Research Act)

The Federal Act on Research involving Human Beings (Human Research Act, HRA) entered into force in 2014. The HRA was enacted to comply with the main objective of Article 118b of the Swiss Constitution. The purpose of the law is to protect the personal rights of individuals, who participate in a research project. A person participating in a research project must be able to be assured that he or she is not exposing to disproportionate risks[7].

HRA. Chapter 1 – General provisions

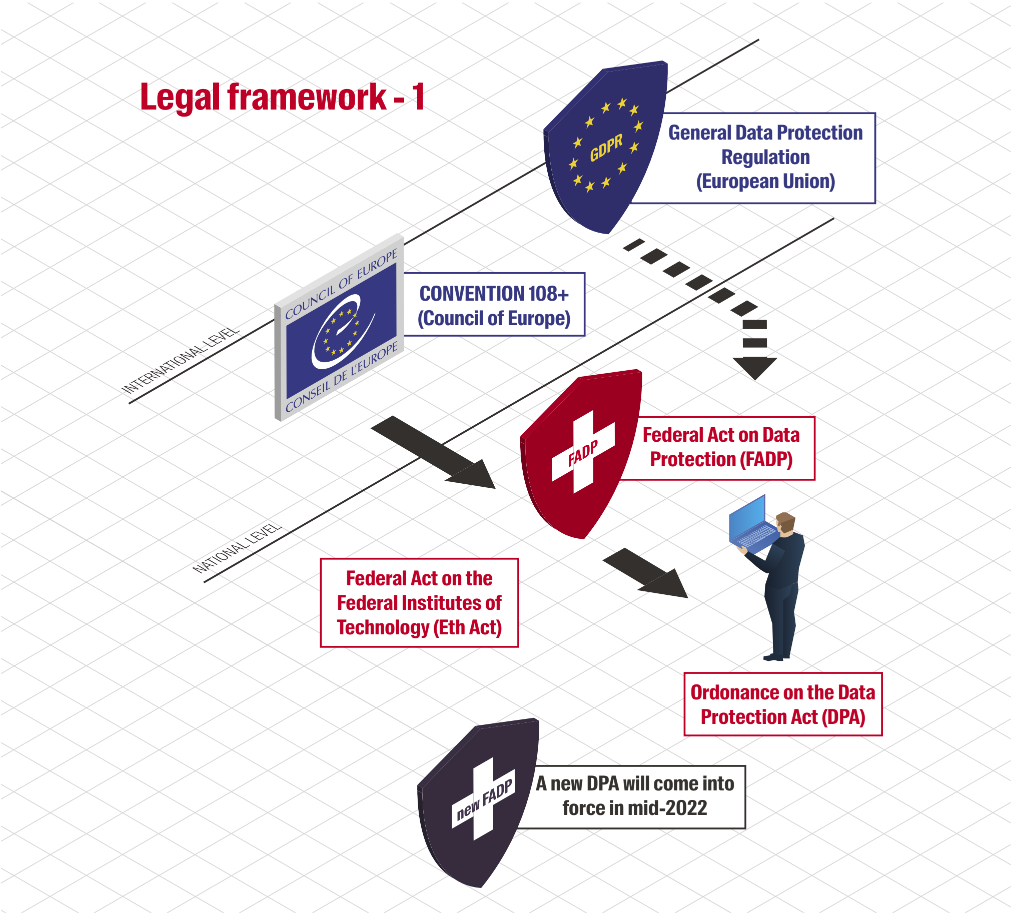
Section 1. Purpose, Scope and Definitions

Art. 1 Purpose

[1] The purpose of this Act is to protect the dignity, privacy and health of human beings involved in research.

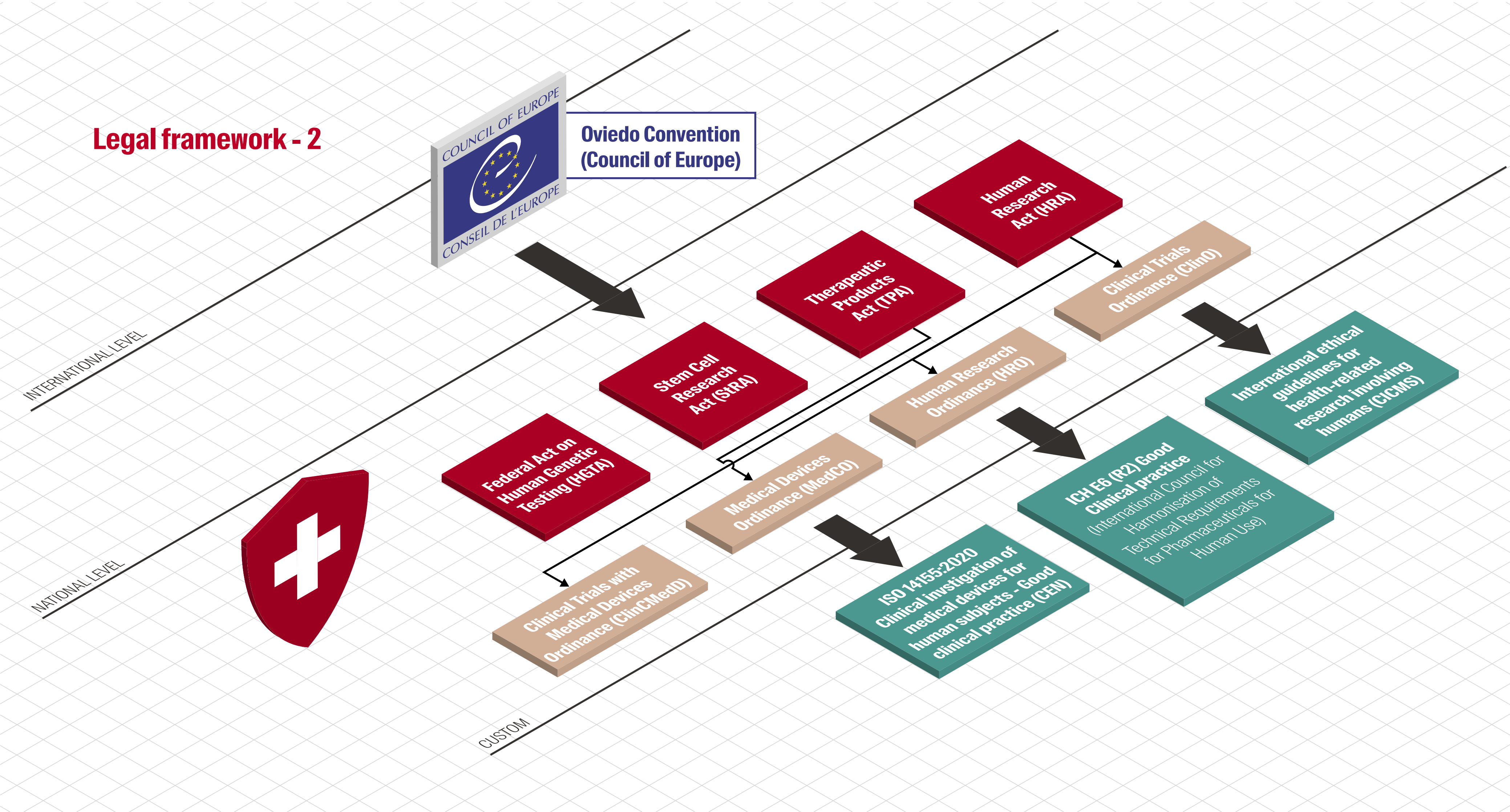
[2] It is also designed to:

- create favourable conditions for research involving human beings;
- help to ensure the quality of research involving human beings;
- ensure the transparency of research involving human beings.


Legal framework - 1
GDPR
General Data Protection Regulation (European Union)
COUNCIL OF EUROPE
CONSEIL DE L'EUROPE
CONVENTION 108+ (Council of Europe)
INTERNATIONAL LEVEL
NATIONAL LEVEL
FADP
Federal Act on Data Protection (FADP)
Federal Act on the Federal Institutes of Technology (Eth Act)
Ordonance on the Data Protection Act (DPA)
new FADP
A new DPA will come into force in mid-2022

[7] Message sur la loi fédérale relative à la recherche sur l'être humain, FF 2009 7259

Legal framework - 2


COUNCIL OF EUROPE
CONSEIL DE L'EUROPE
Oviedo Convention (Council of Europe)
INTERNATIONAL LEVEL
NATIONAL LEVEL
CUSTOM
Human Research Act (HRA)
Therapeutic Products Act (TPA)
Stem Cell Research Act (StRA)
Federal Act on Human Genetic Testing (HGTA)
Clinical Trials Ordinance (ClinO)
Human Research Ordinance (HRO)
Medical Devices Ordinance (MedCO)
Clinical Trials with Medical Devices Ordinance (ClinCMedD)
International ethical guidelines for health-related research involving humans (CICMS)
ICH E6 (R2) Good Clinical practice (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use)
ISO 14155:2020 Clinical investigation of medical devices for human subjects - Good clinical practice (CEN)

4.3 EU-General Data Protection Regulation (EU-GDPR)

In an effort to update the protection of personal data in the European Union, the European Parliament has adopted the EU-General Data Protection Regulation (EU-GDPR) which entered into force 2016 and is applicable since May 2018 in the EEA. This is a new, more restrictive version of the law and it supersedes Directive 95/46/EC, adopted in 1995. This revision of the law was motivated by the rapid evolution of technologies, whose corollary is the exponential sharing of personal data.

Switzerland is not a member of the EEA, so this Regulation does not apply in Switzerland, but is an obvious source of inspiration. Switzerland has been more observant in its approach and has waited to know more about the European developments before aligning itself. Switzerland signed the Council of Europe's' revised convention on data protection known as «Convention 108+» which will probably enter into force in the coming 1-2 years[8].

Even though the EU-GDPR generally does not apply in Switzerland, this regulation has an extraterritorial reach and can apply in some cases to research conducted by EPFL researchers.

- The first situation is when goods or services are offered to people in the EEA (for example, developing and making available in EEA countries an application in the App Store).
- The second situation is when people's behaviour within the EEA territory is monitored. For example, tracking on the Internet, tracking via wearable or other smart devices, observation of geo-localization, analyzation of the use of websites, behavioural studies, etc.
- The third situation relates to the processing of data in the context of the activities of a partner established in the EEA. For example, EPFL processes personal data in the context of the research project of a partner based in the EEA.

[8] The requirements of the Convention 108+ have already been included in the revised FADP.

4.4. Legal basis

As a federal entity, EPFL shall process data only if there is a statutory basis for doing so (art. 34 FADP). Researchers need to ensure that they have a legal basis for the collection and processing of personal data for research.

Benefiting from a formal legal basis, the Federal Institutes of Technology (ETH) will rely mainly on art. 36c of the Federal Act on the Federal Institutes of Technology (ETH Act). This legal basis allows EPFL to process personal data, sensitive personal data and personality profiles in the context of research projects, provided that they are necessary for the research project in question (art. 36c para. 1 ETH Act). In addition, the FADP is applicable (art. 36c para. 2 ETH Act), which means that the general law will always apply if the specific law does not provide an answer. In this case, the ETH Act only formulates few obligations, which means that the general law - the FADP- will largely apply.

In the field of research, a distinction must be made between the general legal framework applicable to any processing of personal data and the specific framework that applies when researchers process health data in connection with a study or clinical trial (for instance).

ETH Act - Chapter 6. Data processing

Section 2. Handling of personal data in research projects

Art. 36c Data processing

- Within the scope of research projects, the two federal institutes of technology and the four research institutes within the ETH Domain may process personal data, including particularly sensitive personal data and personal profiles, insofar as this is required for the given research project.
- In doing so, they shall ensure compliance with the provisions of the Federal Act of 19 June 1992 on Data Protection.

4.5. Sensitive data

Note that the so-called sensitive personal data is subject to a different and more protective regime. The consent requirements are higher (express consent is required) and the requirements for the content of the information given to the participant are stricter in general. The requirements may be even more stringent if health data are collected or used in the context of research on human diseases or on the structure and functioning of the human body (art. 2 al. 1 HRA).

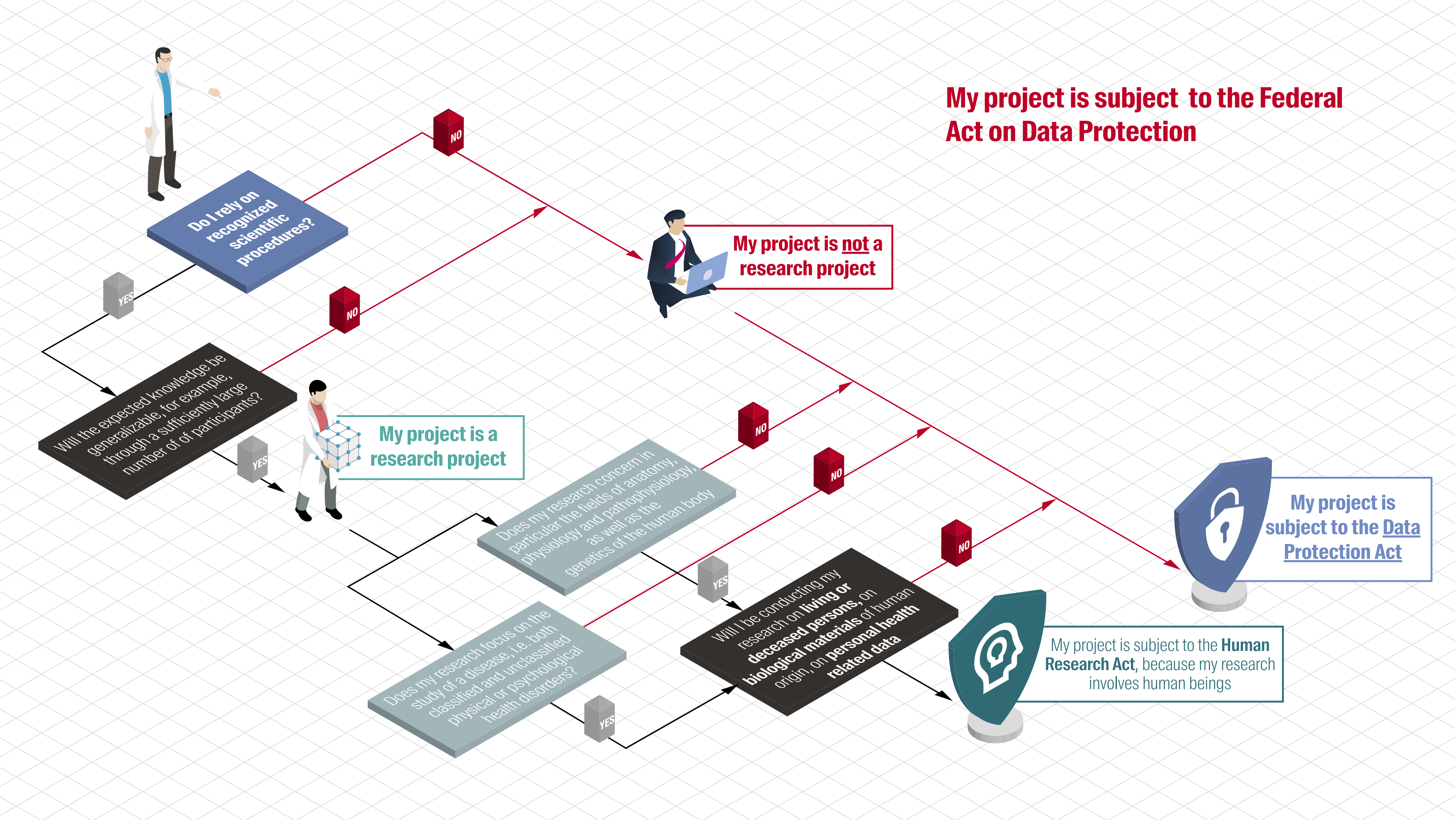
If the purpose of the project is about finding out more about a disease or the function/structure of the human body, the HRA applies (art. 2 HRA). For any other projects involving health-related data, the FADP (and ETH Act) apply. (See figure below.)

The ground of legitimacy differs depends on whether the HRA or the FDPA is applicable. Under the FDPA, EPFL may process personal data on the basis of the ETH Act, even when health data is involved.

Under the HRA, on the other hand, the consent of the participants in a clinical study is in principle required (art. 7 HRA). In addition, the re-use of personal health-related data for human research purposes requires the authorization of the competent ethics commission (Art. 45 para. 1 HRA cum Art. 33 HRO).

Cantonal ethics committee and EPFL Human Research Ethics Committee (HREC)

The mandate of the EPFL Human Research Ethics Committee is to evaluate projects from EPFL researchers that involve human participants and/or personal data. These projects do not fall within the scope of the Federal Act on Research involving Human Beings (Human Research Act, HRA). The projects within the scope of the HRA must be evaluated by the cantonal ethics committee.



My project is subject to the Federal Act on Data Protection

Do I rely on recognized scientific procedures?

Will the expected knowledge be generalizable, for example, through a sufficiently large number of of participants?

My project is a research project

Does my research concern in particular the fields of anatomy, physiology and pathophysiology, as well as the genetics of the human body

Does my research focus on the study of a disease, i.e. both classified and unclassified physical or psychological health disorders?

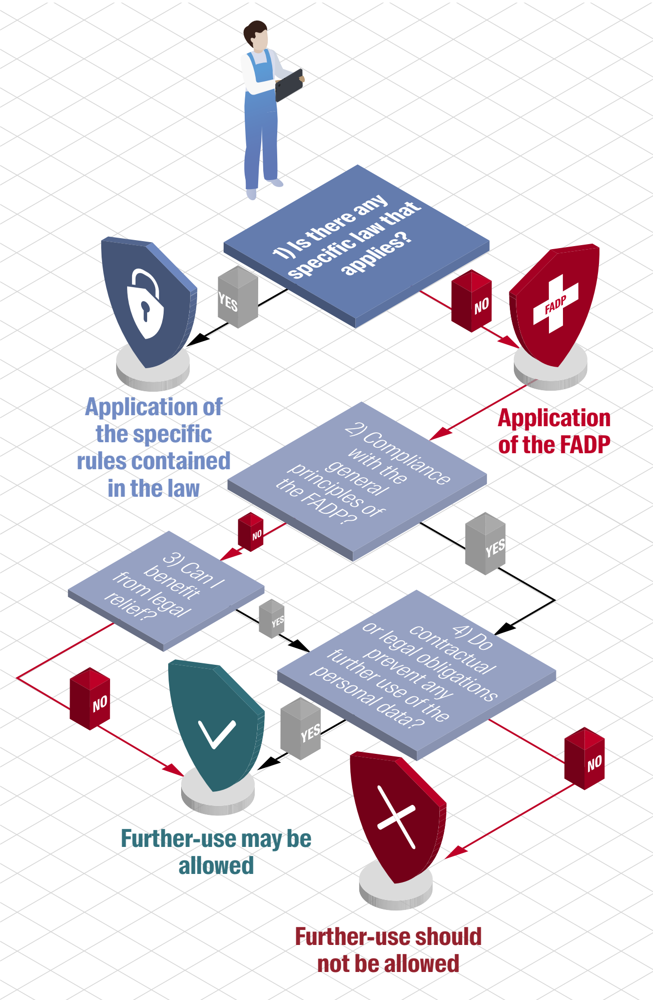
Will I be conducting my research on **living or deceased persons,** on **biological materials** of human origin, on **personal health related data**

My project is not a research project

My project is subject to the Data Protection Act

My project is subject to the **Human Research Act**, because my research involves human beings

4.6. Re-use of data


1) Is there any specific law that applies?
YES
NO
FADP
Application of the specific rules contained in the law
Application of the FADP
2) Compliance with the general principles of the FADP?
NO
YES
3) Can I benefit from legal relief?
YES
NO
4) Do contractual or legal obligations prevent any further use of the personal data?
YES
NO
Further-use may be allowed
Further-use should not be allowed

When the HRA is not applicable, which is frequently the case at EPFL, the general rules of the FADP apply for any re-use of personal data.

According to David Rosenthal[9], the admissibility of using personal data already collected for other purposes must be examined in different steps.

- The first step is to investigate whether there is any special regulation that contains specific rules for the area in question, in which case these will take precedence (such as the rules contained in the HRA).
- The second step is to examine whether the further processing of personal data complies with the general principles of data protection. As a reminder, the general principles include lawfulness, transparency, good faith, purpose limitation, proportionality, accuracy and security. In the context of further use of personal data, the principle of purpose limitation is of particular importance, as the further use must generally be compatible with the purpose originally communicated to the individuals.
- Thirdly, the research benefits from a substantial relief if the purpose of the processing of personal data does not concern the identity of individuals (such as a research purpose), but the conditions of Art. 39 FADP must be complied with[10]. For example, the secondary purpose does not need to be recognizable to the individuals, and the communication of the personal data is more flexible.
- A final step is to consider whether contractual or legal obligations prevent any further use of the personal data (contractual promises or commitments, confidentiality clauses, respect of the official secrecy) or whether further use is inappropriate because it could cause damage to the image or provoke negative reactions among the population[11].

It should be noted that the reasoning in the context of the GDPR is similar but is based on the criteria provided for in Article 5, Article 6 par. 4, Article 13 par. 4, Art. 14 par. 5, Art. 89 par. 1 GDPR, and does not provide for the same reliefs as the FADP (it is actually less flexible).

New FDPA - Chapter 6. Special provisions for the processing of personal data by federal bodies

Art. 39 Processing for non-personal purposes

[1] Federal bodies may process personal data for purposes that do not relate to persons, in particular in the context of non-personal purposes, in particular for research, planning or statistical purposes, if the following conditions are met:

- the data is rendered anonymous as soon as the purpose of the processing permits;
- the federal body communicates sensitive data to private entities only in a form which does not allow the data subjects to be identified;
- the recipient shall only disclose the data to third parties with the consent of the federal body that transmitted the data to him;
- the results of the processing are published only in a form that does not allow the data subjects to be identified.

[2] Articles 6 paragraph 3, 34 paragraph 2 and 36 paragraph 1 do not apply.

[9] ROSENTHAL, Die rechtlichen und gefühlten Grenzen der Zweitnutzung von Personendaten, Forum, in sic!4|2021

[10] In particular, the data shall be anonymized as soon as possible.

[11] David ROSENTHAL, Die rechtlichen und gefühlten Grenzen der Zweitnutzung von Personendaten, Forum, in sic!4|2021

4.7. In the context of the HRA

When collecting data, researchers should determine from the outset whether they intend to re-use the data for other research purposes. If so, the consent of the participants for the 'further use' of the data may need to be obtained at the time of collection (art. 17 HRA). Furthermore, in the context of the HRA, the re-use of personal data in a new research project requires the approval by the competent ethics commission (art. 45 par. 1 HRA cum art. 33 HRO).

Despite obtaining the consent for the further-use of personal data is the rule, the legislator provided for an exception when it is it is impossible or disproportionately difficult to obtain consent (i), when there is no documented refusal of further-use (ii), and if the interest of the research outweighs the interest of the participant (iii) (art. 34 HRA). Those conditions are cumulative. IIn this case, the approval of the competent ethics commission is always required[12]. Some ethics commissions generally consider that the inclusion of less than 100 datasets/samples, and that the availability of a broad consent for recent data is not disproportionate difficulties to ask for consent[13].

Since the Human Research Act applies a more flexible regime for anonymous data (like the FADP), the anonymization of personal data may appear to be a solution that allows researchers to use (or reuse) these data without being forced to follow an excessively cumbersome procedure[14]. However, because researchers often do not know if/for what purpose they will use the data in the future, it is often difficult to know whether it will be necessary to identify these individuals at a later stage. Furthermore, a cautious approach is advised in view of the approach chosen by the legislator in the HRA[15]. This is why in practice it is preferable to proceed to a reversible encryption of these data, by coding them[16]. Furthermore, anonymization de facto deprives the persons concerned of the results of the research, which is rarely ideal[17].

The HRA establishes a complex system[18] of disclosures, categorizations, and consent requirements. Apart from the procedure of art. 34 HRA, a specific consent or a broad consent is generally required for the further-use of personal data.

In summary, the HRA provides for the following[19]:

Link with the individual	Genetic personal data	Non-genetic personal data
Raw data	Information to participant + Specific consent for each use in a particular research project	Information to participant + Broad consent for 'research purposes'
Pseudonymized data	Information to participant + Broad consent for 'research purposes'	Information to participant + right to withdraw
Anonymized data	Information to participant + right to withdraw (at the moment of collection)	HRA and FADP do not apply.

[12] DRIESSEN Susanne / CHRISTEN Andri / GERVASONI Pietro, Humanforschung, Weiterverwendung und informierte Einwilligung, in: Jusletter 1. Februar 2021, §52

[13] DRIESSEN Susanne / CHRISTEN Andri / GERVASONI Pietro, Humanforschung, Weiterverwendung und informierte Einwilligung, in: Jusletter 1. Februar 2021, §52

[14] ISSENHUTH-SCHARLY GHISLAINE, Autonomie individuelle et biobanques, Etude de droit comparé, Schulthess (Editions romandes), 2009, p. 210

[15] ERARD Frédéric, Les données codées dans le contexte de la recherche: personnelles ou anonymes?, in AJP/PJA 5/2021, p. 606ss

[16] ISSENHUTH-SCHARLY GHISLAINE, Autonomie individuelle et biobanques, Etude de droit comparé, Schulthess (Editions romandes), 2009, p. 214

[17] Ibidem, p. 215

[18] JUNOD Valérie / ELGER Bernice, Données codées, non-codées ou anonymes : in: Jusletter 10. Dezember 2018, N 16ss

[19] Repris de l'article de SALATHE Michelle et DRIESSEN Susanne, Consentement général : un modèle uniforme pour faciliter la recherche sur tout le territoire suisse, in Bulletin de l'ASSM, 03/2016

#5 Research Personal Data Life Cycle

Before starting your research project, it is important to ensure that you have a legal basis or the consent of the participants, as mentioned in the previous section.

As a researcher, and especially if you are funded by a granting agency, you must write a Data management plan (DMP) before starting your project.

A DMP describes the steps to be taken at the different stages of the research cycle to ensure successful data curation and preservation. There are several stages in the research data life cycle, e.g. data creation, data processing, data analysis, etc.

5.1. Goal of a DMP

The main goal of a DMP is :

- to plan your data strategy ahead of your project in order
- to anticipate your needs in terms of
  - *resources (servers, hard drives, data curation and preservation, softwares, etc.) and*
  - *good practices (standardized documentation, metadata collection, naming convention, data security, open science, regular backup, etc.).*

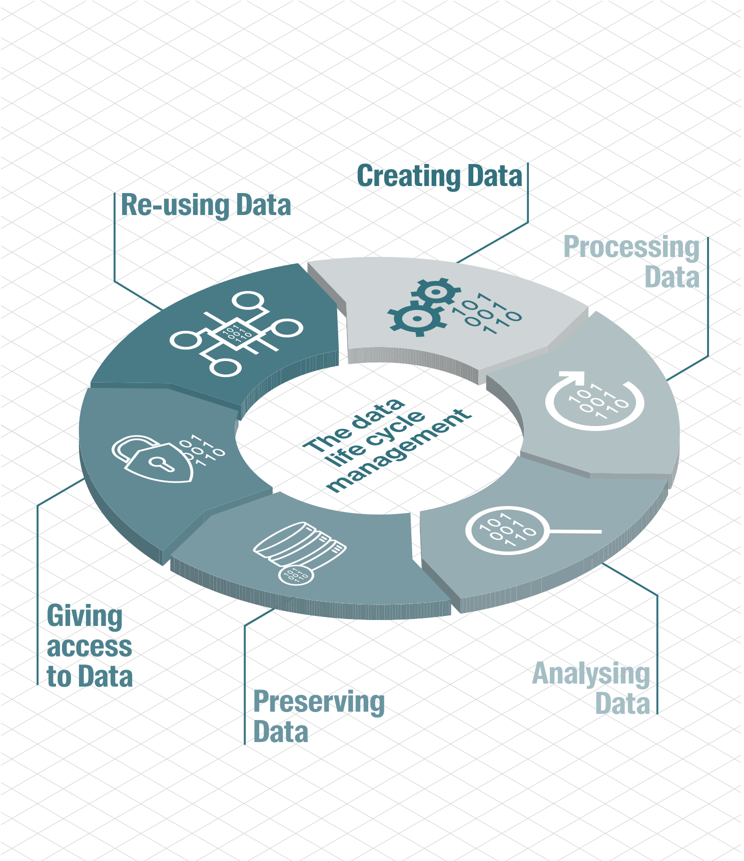
Standardized procedures enable any collaborator to understand your processes.

The DMP is also a very useful tool from a data protection perspective.

Indeed, the clarity that this document provides in terms of data life cycle management and the measures - technical and organizational - that are taken to protect the data, helps the researcher to manage the data effectively and to reduce the risk of data loss or other threats that could make the data unreadable or unusable.

Note that a DMP is often requested by the funding body. Each funding body might have its own template/requirements. The Library can provide you the necessary support.

5.2. The data life cycle management



Data Collection, data processing and analyzing

Duty to inform

From the research point of view, you may distinguish two use cases: (A) Data collected directly from data subjects, (B) Indirectly collected data.

Examples:

A. Data from a survey, questionnaires, interviews, ...

B. Data from public statistics, surveys from polling institutes, data from administrative data bases, data extracted from the web, ...

From the legal perspective, please take in mind this general principle:

The data controller shall adequately inform the data subject of the collection of personal data, whether or not it is collected from the data subject.

In particular, the data subject must be informed of:

- the name of the data controller
- the concerned categories of data
- the purpose of the processing (and “secondary purpose” or further use of data, if applicable)
- the duration of data retention
- the procedure of exercising the data subject associated rights
- the name of the Data Protection Officer
- the name of the State or international organization to which data is transferred (if applicable)

Within the FADP and GDPR, the so-called **«research privilege»** makes it possible to use personal data collected in a previous research project or provided by a public body if it is anonymized or pseudonymized. No further justification (such as consent) is then necessary for data processing without personal reference.

→ Note that, the duty to inform is a notion that applies not only at the time of collection but throughout the data life cycle in the event of a change in the purposes of the project, the categories of data collected, the legal basis, the recipients, etc.

Ref. Art. 19 FADP (Art. 13&14 GDPR), 39 FADP (Art. 89 GDPR)

Please also take in mind:

- **Pseudonymization:** as seen in the main definitions section, the pseudonymization is a valuable security measure to reduce risks for data subject. You should implement this measure since the data collection.
- **Data Processors:** : If a service provider is employed, a contract determines the obligations and commitments of each party. The data controller must communicate all instructions on the data processing to the data processor and respect the principles of the data protection regulations. In particular, include a confidentiality agreement and ensure the security of the information systems used.

Data minimization

Should I need all the data I planned to collect and process? Collect and process data only when adequate, relevant and not excessive in relation to the purposes

Example:

A PI conducts a research project whose aim is to study the negative effects of a particular protein in diabetic patients. She sends participants a general questionnaire, which includes specific questions about their sexual behavior. It would be irrelevant and excessive to obtain such information from an individual who was participating to a diabetic study.

Ref. Art. 6 FADP, Art. 5 GDPR

Privacy by design and by default

As recalled by the FDPIC[20], the FADP enshrines the principles of privacy by design (data protection through technology design) and privacy by default (only data that is absolutely necessary to a specific purpose is processed, and this should be set out before data processing starts).

These principles require authorities and businesses to implement the processing principles of the FADP from the planning stage by putting in place appropriate technical and organizational measures.

Privacy by design requires that their applications and similar are designed in such a way that data is anonymized or deleted by default.

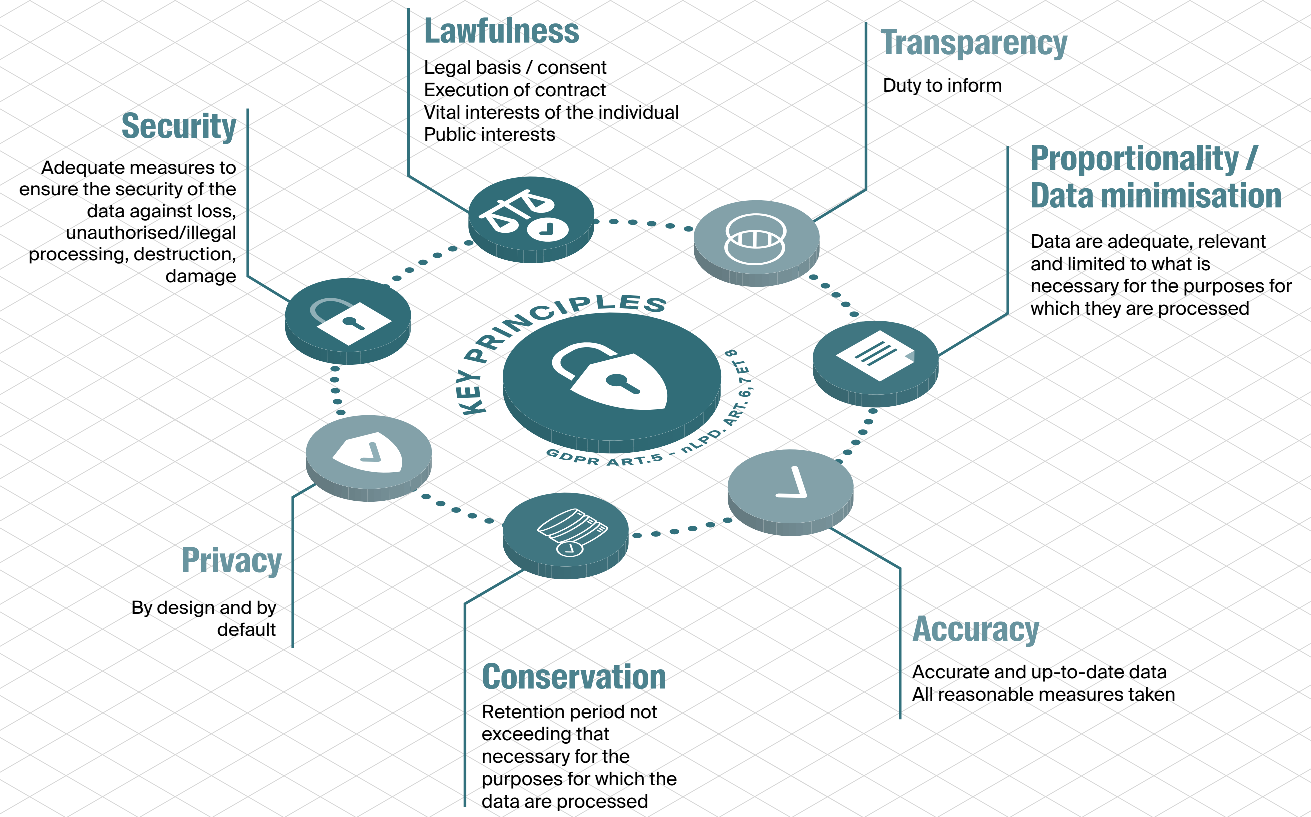
Privacy by default protects users of private online offerings who have not look into the terms of use or the associated right of objection as only the data that is absolutely necessary for the intended purpose is processed, as long as users do not take action and allow further processing.

Example:

The development of SwissCovid and NotifyMe apps are valuable examples of the application of privacy by design principle

Ref. Art. 7 FADP, Art. 25 GDPR

Key principles



Lawfulness

Legal basis / consent
Execution of contract
Vital interests of the individual
Public interests

Transparency

Duty to inform

Security

Adequate measures to ensure the security of the data against loss, unauthorised/illegal processing, destruction, damage

Proportionality / Data minimisation

Data are adequate, relevant and limited to what is necessary for the purposes for which they are processed

Privacy

By design and by default

Accuracy

Accurate and up-to-date data
All reasonable measures taken

Conservation

Retention period not exceeding that necessary for the purposes for which the data are processed

[20] The new DP Act from the FDPIC's perspective, February 2021

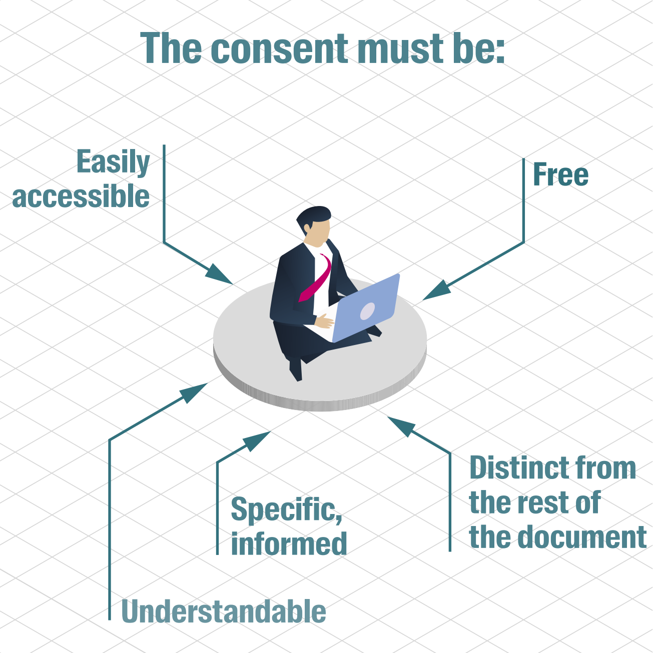
Accuracy

Anyone processing personal data must ensure that it is accurate. He/she takes all appropriate measures to rectify, erase or destroy inaccurate or incomplete data that are inaccurate or incomplete with regard to the purposes for which they are collected processed. The appropriateness of the measure depends in particular on the type and extent of the processing and the risk that the processing of the data in data processing in question poses to the personality or fundamental rights of the data subjects.

Ref Art. 6 FADP, Art. 5 GDPR

Consent

When you collect data, you usually ask data subjects for consent. A comprehensive informed consent is a crucial requirement in research.



The consent must be:

Easily accessible

Free

Specific, informed

Distinct from the rest of the document

Understandable

Consent for sensitive/health-related data

For research falling under the responsibility of swissethics there are specific templates for consent available here.

Data protection impact assessment (DPIA)

- Data Controllers are obliged to carry out a data protection impact assessment (DPIA) on any new project likely to involve a high risk to the personality or fundamental rights of the data subject.
- As recalled by the FDPIC, the high risk comes from the nature, scope, context and purposes of processing – particularly when using new technologies. In particular, processing is deemed high risk if profiling or extensive processing of sensitive data is planned.
- The CNIL, the French supervisory DP authority provided a lot of documentation to understand when abd how a DPIA must be done. Please refer here.
- DPIAs must be prepared before beginning any data processing activity.
- Controllers may abstain from establishing a DPIA, for example, if they use a system, product or service that is certified for the intended use by a recognized independent certification organization
- When you submit a project to the Ethical Committee, a first data protection assessment is provided by the DAR Legal counsels. In case of high risks for natural persons, the DPO must be involved and can accompany the researchers in carrying out a DPIA
- DPIA tool: The CNIL has developed an open source tool that can be useful to carry out a DPIA : PIA tool.

Ref. Art. 22 FADP (Art. 35 GDPR)

Security

Security is a set of organizational and technical measures to protect the confidentiality, integrity, availability and traceability of data to enable the organization to meet its objectives and obligations. Security is not an end in itself, but an indispensable means. It is planned and the means implemented are proportional to the value of the assets to be protected and/or the legal, regulatory or contractual obligations.

The guide provides a set of measures to be implemented in the appendix. These should be planned prior to data collection and processing as they apply to the environment in which the data will be processed. The implementation of these measures, where applicable, should be required to partners or subcontractors. The contact points are the Heads of IT and the AdminIT.

IT Security is an important requirement for complying with data protection laws. Organisational measures and physical security are important elements of IT security and should not be neglected (e.g. against theft of mobile storage media or IT equipment, or access to password lists).

Ref. Art. 7 and 8 nFDAP, LEX 6.5.1, LEX 6.1.4 and LEX 6.1.3 (art. 32 GDPR)

Example:

An IP runs a project where it is necessary to ensure that only authorised people can change data and that changes are tracked. Access rights management is in place as well as logging of processing. The logs are also protected by access rights.

The following questions are an aid to completing a Data Management Plan. The security measures in the appendix help to answer them.

Data Sharing

- Do you share personal data with another EPFL entity? It is important to comply with the need to know principle. If you work with another internal unit, please verify who must have access data.
- Do you share/transfer personal research data with anyone outside of EPFL ? (EU/non EU/USA) We must comply with the legal framework. For instance, if you work with a subcontractor based in the USA, you must put in place supplementary measures or safeguards (e.g. contractual, IT security) in order to ensure data is protected. Some technical safeguards that can be implemented: encryption, pseudonymization, honest broker...
- Do you use public cloud solutions? If yes, do you assess the level of risk (legal, reglementary and contractual risk) prior to save data into the Cloud?

Output

- What is the output from a data point of view (scoring, decision, personal advice...)?
- Does the data processing create automated decisions impacting people? Remember that this kind of data processing can represent a high risk for the individual. This is a situation demanding a DPIA before starting your project
- Do you anonymize data for publishing purpose?

Data retention period, Archiving, Data deletion

- How long do you retain data?
- Do plan to archive data?
- Have you defined the archiving process?
- Do you plan to delete data?
- Have you defined the deletion process?

Training

- Is your staff aware of data protection requirements that covers personally identifiable information (PI)? If not, please ask the DPO

Do you organize a training of employees on data protection in your laboratory? If not, please ask the DPO.

Organization and IT security

- Which technical and organizational measures do you take? (e.g. for technical measures: encryption, encryption key management process, ...)
- How do you ensure that they are appropriate to the risks? e.g. tests, audits,...
- Is a process in place to review event logs?
- Do you implement Privacy-by-Design or by-Default?
- Do you maintain security configuration standards for information systems and applications?
- How do you ensure the data traceability?
- Do you back up data?
- Is the data processed in a manner that ensures

appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage, using appropriate technical or organisational measures?

- Is this infrastructure in place based on agreed specifications with DSI?

Asset control

- Do you have a process in place for granting and revoking appropriate user access (incl. privileged access rights and roles)?

- Do you regularly review users access to ensure only needed privileges are applied?

- Do you have standards for isolating sensitive data and procedures and technologies in place to protect it from unauthorized access and tampering?

- Have you defined rules and processes to manage access to IT resources for guests and external users?

Preserving

- Would you like to preserve the datasets underlying your publications?
- Are you leaving your lab and have data or code to archive?
- Would you like to get expert support for data curation?
- Do you have reference datasets that you would like to preserve?

Contact acoua@epfl.ch, if you want professional support to:

- Define the specific curation needs for your datasets and related code
- Help you in appraising the datasets and related code to be preserved
- Ingest your dataset(s) and related code in a trustworthy, safe and EPFL-backed environment
- Periodic control of the integrity of the datasets via audits and data curation actions
- Provision of reports on the state of the preserved dataset to the research group

Reuse

- The re-use of data in research is quite common. From the data protection perspective, it is important to verify if you have previously obtained a consent from the data subject to re-use his or her data, or if you have another legal basis or contractual obligation that allow you to process his or her personal data. You can re-use, without asking a new consent, anonymized data.

- If you are using data from social media networks you must also ensure that your intended use of the data complies with any terms and conditions published by the network[21].

5.3 Ethical authorizations

Research projects that involve human beings and/or personal data need to obtain an ethics authorization. If a research project falls within the scope of the Federal Act on Research involving Human Beings, LRH (or: Human Research Act, HRA), it needs to be submitted for review to the cantonal ethics committee. If a project involves human subjects and/or personal data but does not fall within the scope of the Human Research Act, it needs to be evaluated by the EPFL institutional ethics committee, the EPFL Human Research Ethics Committee (HREC).

For guidance which ethics committee is competent and how to submit the project, please contact the Research Office - Ethics Affairs at research@epfl.ch.

References: https://www.assm.ch/en/Publications/Handbooks.html

[21] This apply generally to any data obtained from a third-party.

#6 Health Data Management

Health data includes, but is not limited to:

- Patient demographics
- Medical notes
- Laboratory test results
- Procedures (SOP) and surgeries
- Imaging, like x-rays, computerized tomography (CT), and MRI
- Prescriptions
- Referrals and other communication
- Provider information, etc.

Health data belong to sensitive data and must be protected in an appropriate way.

The lack of an adequate protection entails significant legal, financial, reputational and scientific risks. For example, there can be criminal and civil sanctions if a research project that involves sensitive data processing is not compliant (for example no ethics authorization was obtained, a clinical trial requirement is not met and participants encounter a health problem).

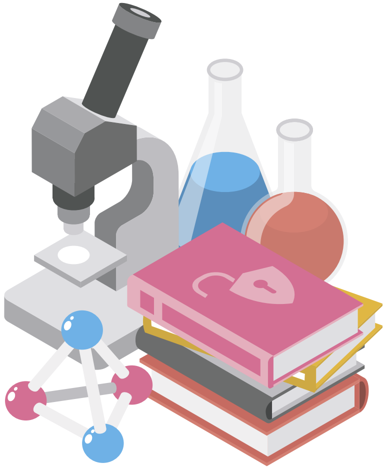
IT and physical security measures to implement are described in the Appendix. They include security measures to protect personal data and some additional measures.

Different IT systems exist at EPFL (e.g. RedCap) that are compliant with the strict regulations on health data.

The SV-IT team can help you to implement such systems and provide the adequate support in case of need.

#7 Open Science and Data Protection

This section aims to streamline the balancing act between principles of open science and data protection of personal data. It aims to provide a framework for related obligations and possible opportunities for research practice.



7.1. Open Research Data

Open research data (ORD) is research data that is publicly available for re-use, in an appropriate form and with minimal restrictions, e.g. permissive licenses. The principle of openness is not in contradiction to authorship or possible commercial exploitation of such data. The same applies to the presence of potentially personal data in ORD.

In Switzerland, ORD is becoming the preferred default option in any research practice. In some instances, such as receiving grants from certain funders or working for certain institutions, a stronger ORD mandate is imposed.

Personal data is an important parameter to consider in ORD. Collecting or processing personal data is not in most cases a barrier to ORD altogether. Instead, appropriate ORD practices that take into account the nature of the research and the type of personal data must be implemented to accommodate data protection needs.

There are many clear benefits of ORD, such as increased accessibility to research outcomes, higher visibility and potentially higher impact of research, and deeper and larger collaboration prospects.

7.2 Opportunities of ORD and Data Protection

The presence of personal data within the research is an opportunity that can bring benefits to your research, your collaboration, and your research team. Understanding and following the rules and regulations of data protection can help to build a better understanding of research data. It can also help to focus on data quality and increase trust in data. With better insights, potential hidden data biases can be identified. Opportunities of ORD and obligations of personal data protection are not in conflict. Requirements related to the collection and processing of personal data provide an opportunity to approach your data in a systematic and organised manner. With better planning and preparation, the number of possible complications and dilemmas can decrease dramatically.

7.3. Good Practice for Balancing Data Protection and ORD

As the balancing act of data protection and ORD is becoming a daily practice in scientific research across disciplines, there are many examples of good practice available. Getting familiar with these practices and incorporating them into your everyday research practice will improve your overall research process.

There is a set of questions that a researcher should ask when he/she is processing personal data as a part of a research activity. The obtained information can significantly increase understanding of the data. With enough knowledge of the data, you can limit the presence of personal data to a minimum. Limitations connected to personal data processing can provide you with an opportunity to give preference to data quality and limit data creation/collection to the minimum needed.

Trust is essential for research. By applying all necessary measures connected to personal data processing, one increases understanding of data and naturally gives preference to the quality of data. As an implication, trust in the data increases. Trust in data is built not only on a personal level but also within the research group. Among collaborators, trust in data and data processing is essential.

New research areas like machine learning and artificial intelligence are increasingly providing us with evidence of hidden biases in research data. The balancing act of ORD and processing of personal data in research can help you to identify potential biases early and take necessary measures to mitigate the risks.

7.4. Subject Centric

Be prepared to explain your ORD intentions, plans, and obligations to your data subjects. Build trust with data subjects from the start by making sure that they have a clear understanding of ORD and its implications.

Example:

Adjust the text of an informed consent form to reflect your ORD plans. Consider including information on data publishing and long-term online availability.

7.5. Plan for ORD

Invest the time in preparation and planning. When planning, explore various techniques to minimise the presence of personal data. Analyse your hypothesis and design of experiments. Define what is minimum needed data. Consider the level of detail needed for your research. If you cannot avoid personal data during collection, anonymize, when possible.

Reducing the amount of data, dividing datasets, avoiding unnecessary details, and anonymization must take place as soon as possible during the research process. When applied only to research outcomes, the reproducibility of your research may be compromised.

Separate

Apply a fine-grained vision of your data - consider open/close not at the level of projects, or datasets, but when possible at the level of data fields. Split dataset into more parts, divide personal and not personal data.

Reduce

Reduce the amount of personal data to a minimum. Always check beforehand if the data that you wish to collect is necessary for your research. Ask control questions, such as “Do I need this information for my research?”.

Example:

Planned data collection: Name, Surname, Sex, Address, Age, Telephone number, E-mail address

Minimum needed data collection: Initials, Surname, Sex, City, E-mail address

Avoid

Avoid collection of exact data, if possible. Lessen the collection of explicit data. Give preference to clustering, e.g. instead

Example:

Address (street, city, country) -> Area (city, country)

Age (21, 34, 46, etc.) -> Age group (20-30, 30-40, etc.)

Anonymize

There are various techniques of anonymization. See chapter on Anonymization for more details. When you anonymize data that you have created, you manage the anonymization process. However, when you receive data from a data provider, make sure that the data is anonymized. The best way to do that is to seek information on the anonymization process directly from the data provider.

7.6. As Open as Possible

Always apply the principle of openness. Limit closed data to a minimum. If data have to be closed, make sure that the reasons for that are valid. In case the closed data is a necessary approach, you can consider opening the metadata or providing a managed access.

Open Metadata

You always can make your metadata available in an open manner. Plan and prepare for the collection of metadata. Collect it in a moment of creation. Accompanied by a detailed description of data and related instructions, open metadata is a practical approach to balance ORD and data protection, especially for sensitive personal data. Do not forget to be transparent about reasons for data unavailability and potential means of access for legitimated purposes.

Managed Access

If ORD is not an option you can consider providing managed access to trusted users. Plan for this approach early on in the research project. You can plan for a platform and/ or a process that will allow access to data in a controlled way by using secure querying and/or privacy-preserving computation services as an alternative to closing data.

Example:

Swiss personalized Health network

7.7. Available Support

When developing ORD infrastructure, processes and guidelines in your laboratory take into consideration people. Possible fluctuations of researchers within the laboratory can compromise the continuity of the research process. Make sure that continuity is well incorporated into your ORD strategy.

The expertise of others can be extremely beneficial. Do not hesitate to seek external advice and help. When looking for ways to approach your open questions regarding data protection and ORD, explore your research community standards. Engage with the community and inquire about discipline-specific practices. Seek the support of various central services within the institution. Every school offers different research support. Explore possibilities and seek advice. See chapter on Stakeholders and Services for more details.

Example:

Internal on/off-boarding guidance for a laboratory member.

7.8 Anonymization techniques

Anonymization, also known as de-identification, is the process of modifying personal data in such a way that they are no longer, or only with disproportionate effort, that can be correlated with the persons concerned[22]. This is a combination of techniques to be implemented depending on the nature of the data and the possibilities of re-identifying the persons whose data are processed:

- Removal of direct identifiers (e.g. account number, insurance card)
- Suppression, generalization (changing a ZIP code by a state code e.g.), perturbation (replacing some values randomly), permutation (exchanging values between records), sub-sampling (releasing a subset of data), aggregation of quasi-identifiers[23] (e.g. ZIP, birthday, sex, size)

As re-identification techniques increase rapidly, it is important to implement anonymization measures carefully, and to ask the following 3 questions[24] to verify the effectiveness of the chosen measures:

- is it still possible to single out an individual,
- is it still possible to link records relating to an individual, and
- can information be inferred concerning an individual?

Ref. Art. 31 al. 2 let. e FADP

[22] Translated from https://www.edoeb.admin.ch/edoeb/fr/home/protection-des-donnees/generalites/glossaire.html

[23] National Institute of Standards and Technology, NISTIR 8053 - De-Identification of Personal Information (https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf)

[24] Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, p.3 (https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf)

#8 Risk Management: an opportunity for the Research

The Internal Control and Risk Management Department (CIGR) coordinates risk analysis, monitoring and implementation of mitigation measures to ensure the protection of EPFL's value components. These include, among other things, compliance with external and internal legal framework, IT security, human capital, reputation, infrastructure, as well as tangible and intangible financial resources. The CIGR is responsible for monitoring the internal control system, which ensures that the School's activities are in line with the legal basis to which it is subject (compliance).

The purpose of this chapter is to make researchers aware of the importance of complying with data protection rules and not to underestimate the consequences of violating these laws. Non- compliance with data protection laws, such as the Swiss Federal Act on Data Protection and the European Union's General Data Protection Regulation, can directly impact researchers through individual sanctions.

The three major risks related to data protection are the following:

1. **Reputation risk:** Weaknesses in the security of personal data, a poorly defined procedures, and employees with a lack of awareness and training can result in damage to EPFL's image and reputation.

2. **Legal risk:** Sanctions, remedies are possible in case of violation of the data protection law

3. **Financial risk:** The GDPR and FADP provide for financial penalties in case of non-compliance. Funders, notably in European Union, could reduce their funding if EPFL is not considered as a reliable partner in terms of data protection.

Example of potential risks to the privacy of research participant involving health information

The data controller undertakes to collect only the data that is strictly necessary and relevant to the objectives of the research. The research protocol must scientifically justify the use of participant data relating to the health. In this example, an explicit consent of the participants is compulsory.

Potential data breaches from the processing are:

- Illegitimate access to data
- Data leakage
- Unwanted data modification

The management of EPFL requires complete, transparent and up-to-date information. The Risk Management Committee (CRM) continues its work of identifying important risks, thereby facilitating timely and regular communication of these risks to the Management and supervisory bodies. This has resulted in a better overview and increased risk transparency.

The CRM has for missions to develop a system of organization and procedures to ensure legal compliance, guarantee the identification of risks and suggest the implementation of appropriate measures for reducing them to an acceptable level.

Insurances

The role of the Internal Control and Risk Management Department (CIGR) is to manage, adapt and optimize EPFL's insurance coverage. He informs, advises and supports units or individuals regarding specific insurance coverage.

CIGR provides valuable advice to the researcher in the event of a need for specific coverage. He is available to help them in analysis and evaluation of files and needs and follow-up of the steps involved in the processing of claims.

#9 Glossary

CIGR	Control and Risk Management Department
CRM	Risk Management Committee
DMP	Data Management Plan
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
DSI	Information System Direction
FADP	Federal Act of Data Protection

FDPIC	Federal Data Protection and Information Commissioner
GDPR	General Data Protection Regulation
HRA	Human Research Act
ORD	Open Research Data
RD	Research Data
ReO	Research Office
TTO	Technology Transfer Office

#10 Appendix

Security measures to implement

Phases	Security Measures to implement and best practices to apply on your research project		Who
Plan + regular basis	Awareness	› Ensure that the Head of Unit knows that he or she is accountable for the Data Protection of the Project › Remind any project team member using EPFL's electronic infrastructure of their responsibilities as described in LEX 6.1.4 Directive on the Use of EPFL Electronic Infrastructure › Remind any project team member using private computer for professional purpose of their responsibilities as described in LEX 6.1.3 Directive on the Use of Private Computer Equipment for Professional Purpose › Ensure that project team members know what to do in the event of a security incident (See Security Incident) **Tools:** › LEX 6.1.4: https://www.epfl.ch/about/overview/wp-content/uploads/2020/01/LEX-6.1.4_EN.pdf › LEX 6.1.3: https://www.epfl.ch/about/overview/wp-content/uploads/2019/09/6.1.3_d_ordinateur_prive_an.pdf	**Project leader**
Plan	Training	› Ensure that the project team members have completed the basic computer security course on the HR training department website	**Head of Unit**
Plan + in case of change	Users management	› Ensure that each individual working on the systems has a personal identifier on the systems › Ensure that each login account is password protected in accordance with the EPFL guidelines **Tool:** https://www.epfl.ch/campus/services/en/it-services/authentication-gaspar/passwords/	**AdminIT**

Phases	Security Measures to implement and best practices to apply on your research project		Who
Plan + in case of change	Users access rights	› Ensure that access rights to data are managed according to the principle of least privilege (Need-to-know and Need-to-do principles) throughout the data life cycle (shared folders, databases, backups, cold data…) and whatever the medium of the data (paper, hard disk, USB stick…) › On user leave remove his access rights › On user role changes modify accordingly his access rights › Check the project team members access rights once a year and keep a record of this audit › Restrict access to privileged accounts (e.g. root, sudo, admin) to the smallest controlled group of users › Minimize the number of processes/services/applications requiring a privileged account › Restrict or disable remote access to privileged accounts › Limit the general access (i.e. non public) to a managed set of users › Minimize the usage of local accounts › If Health data are concerned, ensure that administrative and user account are protected by a strong authentication system. If no strong authentication system is available, ensure the following rules for the password: – 14 characters minimum – not contain any information from your identity record (name, given name, birth date, username, etc…) – not contain words or sequence of words out of the dictionary – include at least one character from each of the lowercase, uppercase, numbers and special character sets – only ASCII character set (no accent) – always a password that has never been used before **Tool :** Unit or Project Sensitive data Access Control Policy template (see appendix)	**Project Leader / adminIT**
Plan	Traceability	› Enable activity logs to track the addition, modification, reading, transmission and deletion of all personal data, sensitive personal data and personality profile data being processed › Ensure that these logs contain information on the origin and nature of the processing, the identity of the person who carried out the processing, the identity of the recipient (if applicable) and when the processing took place › Ensure that these logs are kept for two years separately from the system in which the personal data are processed ; these logs must be accessible only to the persons responsible in charge of data protection compliance control or according to art. 17 LEX 6.1.4 › Ensure that the users' rights do not allow them to modify or delete these logs › Inform users about the implementation of these logs › Keep track of all copies or extracts of Health data › Log access to secure facilities **Tool:** https://www.fedlex.admin.ch/eli/cc/1993/1945_1945_1945/en	**Project Leader / AdminIT**
Plan	Security incident	› Define a security incident procedure › Ensure that the Head of Unit knows the procedure **Tool:** IT Security Incident: What to do? (see appendix)	**Project Leader / AdminIT**
Plan + regular basis	Users computers and mobile IT	› Ensure that PCs and laptops are secured and managed **Tool :** IT Security Baseline – Hardened PCs and Laptops (https://inside.epfl.ch/secure-it/wp-content/uploads/2019/09/IT-Security-Baseline-Hardened-PCs-and-Laptops-V8.pdf)	**Project Leader / AdminIT**

Phases	Security Measures to implement and best practices to apply on your research project		Who
Plan + regular basis	Servers security	› Ensure that security updates are installed immediately › Migrate to a maintained operating system version before the end of the provider's support › Disable unused network services › Restrict incoming connections to those necessary for these services › Whenever possible, restrict outgoing connections to those necessary for this service › Minimize openings in Diode › When possible, disable all Internet connectivity › Ensure that no other external network access is possible › Define on file servers how access rights are inherited inside sub-folders › For file servers, restrict default access to the corresponding data owner › If the data are computed on multiple systems (computing cluster e.g.), ensure that the level of data protection remains the same as defined (access rights, deletion of copies, traceability...) › Use the EPFL Nessus tool to check for vulnerabilities **Tool :** IT Security Baseline – Servers: https://inside.epfl.ch/secure-it/wp-content/uploads/2019/09/IT-Security-Baseline-Servers-V9-20180507.pdf	**AdminIT**
Plan + regular basis	Websites security	› Ensure compliance to EPFL privacy policy › Maintain a link to EPFL privacy policy › Ensure that administration tools are accessed only by authorized administrators › Ensure that critical updates are installed immediately › Ensure that user inputs are sanitized › Deactivate unsecure communication protocols › Ensure that the urls do never contain a password	**AdminIT**
When needed	Homemade applications	› Verify OWASP top ten vulnerabilities › Ensure passwords are not hardcoded nor saved on any platform like GitHub, GitLab, Confluence › Pay particular attention when implementing cryptographic functions › Use approved cryptographic libraries **Tool:** https://owasp.org/www-project-top-ten/	**AdminIT**
Plan + regular basis	Backups	› Ensure a periodical backup of the data › Ensure that the frequency of backups is compatible with allowable data loss in the event of a problem › Check regularly the integrity and accessibility of the data backed up › Ensure that backup media are stored in a secure location accessible only by authorized persons	**AdminIT**
Plan + When needed	Storage	› Do not store health data on portable storage media to avoid the risks of data breaches by theft or accidental loss. If such storage cannot be avoided, encrypt the data, store the data carrier in a safe and store the encryption key securely in another safe. Define access rights to the safes. › Ensure that personal data and health data cannot be accessed by an unauthorized person or indexed by search engines › Ensure that all copies of data are erased if they are no longer needed. › Encrypt health data-at-rest	**Project Leader / AdminIT**
Plan + regular basis	Cold data storage	› Determine who is responsible of authorizing the deletion of cold data › Determine how long to store the cold data deletion log › See Users access rights measures	**Project Leader**

Phases	Security Measures to implement and best practices to apply on your research project		Who
When needed	Cold data storage	› Ensure that cold data deletions are authorized › Keep a log of cold data deletion › Ensure the irreversibly deletion of useless data, regardless of the data storage media › See Users access rights measures	**AdminIT**
Preserving data	Backups	› Ensure that backup media are protected during transfers if they contain sensitive data (encryption if possible, locked container, 2 persons in charge...)	**Project Leader**
When needed	Data carriers to recycle (hard disk, DVD, USB sticks...)	› Ensure that data are irreversibly erased from data carriers when recycled	**AdminIT**
When needed	Third parties	› Ensure that third parties who need access to EPFL infrastructure and are not EPFL members within the meaning of Article 13 of the ETH Act, sign an undertaking to comply with the provisions of LEX 6.1.4 and appropriate additional provisions taking into account their status as third party users (See art. 5 al. 8 LEX 6.1.4) › If a third party needs access to health data, add to the above-mentioned document clauses prohibiting the copying of the data outside the EPFL infrastructure. If these data must be copied outside the EPFL infrastructure, add an obligation to destroy the copies upon request of the head of the EPFL unit concerned, and at the latest on the last day of validity of the above-mentioned document, as well as an obligation to protect the data and to immediately report any leak to the head of the EPFL unit concerned. **Tools:** › inside.epfl.ch/secure-it/en/compliance/ › www.epfl.ch/about/overview/wp-content/uploads/2020/01/LEX-6.1.4_FR.pdf	**Head of Unit / Project Leader**
Plan	Communi-cations	› Encrypt data in transit	**AdminIT**
When needed	Sub-contractors, partners	› Ensure that contracts include security requirements › Ensure that contracts include clauses to recover or destroy data at the end of the contract › Ensure that contracts include an audit clause	**Head of Unit / Project Leader**
When needed	Data transfer	› Ensure that data are encrypted before the transfer › Use a different channel to transmit the encryption key › Ensure that the recipient of the data/key is the correct one before the transfer **Tool:** pegasus.epfl.ch	**Project Leader / AdminIT**
When needed	Data transfer with external storage media	› If an external storage media is used for sensitive data transfer, ensure that media is protected during transfers	**Project Leader / AdminIT**
Plan	Physical security	› Ensure that confidential paper documents and valuable objects or data carriers (hard drives, USB keys, mobile phones, etc.) are not left unattended › Ensure that everyone locks their desk when they leave the room › Where possible, use a security cable to secure computer equipment › Prefer Hardware encrypted USB Keys	**Head of Unit / Project Leader**
Plan	Physical security for physical servers	› Restrict physical access to servers as well as to backup media and IT infrastructure components › Whenever possible, protect access to the BIOS by a non-default password › Whenever possible, turn off PXE boot (Preboot eXecution Environment) and booting from USB disk › Whenever possible, protect IPMI/DRAC access by a non-default password › Whenever possible, only allow booting from authorized devices (internal hard disk, PXE...) › Protect BIOS/boot configuration modification with a non-default password	**AdminIT**
Plan	"Physical" security for virtual servers	› Restrict interfaces redirections (i.e. USB,...) › Ensute that no other systems can access the virtual disk › Secure the remote access to the machine	**AdminIT**

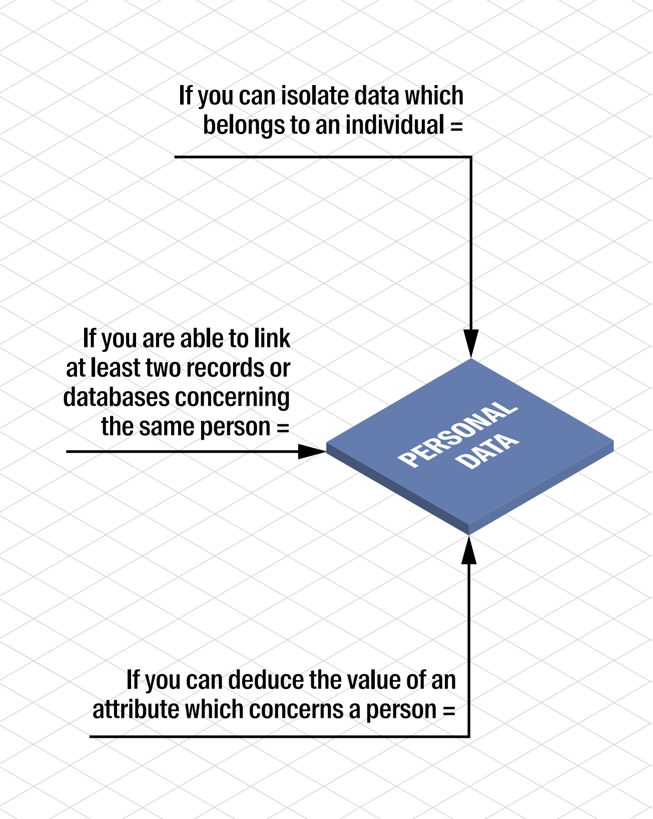
Advanced
Main definitions part 1

Whether or not a person is identifiable is a question that must be asked from the outset. Can the data collected be reasonably traced back to a specific person? If it is possible to associate a piece of information with a person, then this means that the person concerned can be identified, provided that the means of identifying him or her are reasonable in terms of cost and time[25].

A further distinction must be made between whether this information directly identifies the person (such as an identity card), or whether this information indirectly identifies the person by means of additional information (such as the address of a building, which makes it possible to search for the name of the owner of this building)[26] and the identity of the person concerned can therefore be deduced from this additional information.

In other words, the identification can be direct or indirect, depending on how closely or remotely the information relates to the person concerned. The closeness of the link between the information and the individual can only be judged in the individual case, taking into account the financial and time cost factors involved in the identification. In any case, notwithstanding this theoretical distinction, direct or indirect identification falls within the scope of the law, but the distinction is relevant when it comes to coding or pseudonymization of personal data.

→ *Back to 2.2 Data subjects*


If you can isolate data which belongs to an individual =
If you are able to link at least two records or databases concerning the same person =
PERSONAL DATA
If you can deduce the value of an attribute which concerns a person =

[25] Rapport explicatif concernant l'avant-projet de loi fédérale sur la révision totale de la loi sur la protection des données et sur la modification d'autres lois fédérales, Office fédéral de la justice, décembre 2016, p. 42

[26] ATF 138 II consid. 6.1

Advanced
Main definitions part 2

To better distinguish between pseudonymized and anonymized data [vdVE1] , jurisprudence uses the so-called relative theory[27]. This means that we put ourselves in the position of the person who will receive the information/data. In each case it should be determined if this person: (i) is able to directly identify the natural person and (ii) is also willing to make the necessary efforts to identify.

According to the GDPR, which does not apply directly in Switzerland, two conditions must be analyzed to determine whether data is anonymized:

1. is the natural person identifiable?
2. if so, are there objective factors that reasonably prevent the re-identification of the natural person?

In order to answer the first question, the European experts[28] distinguish between whether data can be individualized (individualization), whether they can be linked to each other (correlation), or whether new data can be deduced from the values obtained (inference). If so, whether data can be individualized, correlated, or inferred, in relation to a person, one might consider that one is dealing with personal data. This «absolute» vision has not yet been confirmed by European jurisprudence, which has adopted a relative approach, like the example of Swiss courts. Thus, the Court of Justice of the European Union considers that data is anonymized if the identification of the person concerned is prohibited by law or is practically impossible because it requires a disproportionate effort in terms of time, cost and manpower, so that the risk of identification appears to be insignificant[29].

However, these relative approaches are based on a somewhat outdated view of technological reality. Thus, many legal scholars[30] question this relative theory on the grounds that Big Data and technological advances make attempts at anonymization ineffective. Therefore, the so-called absolute theory, which focuses on the theoretical possibility of re-identification[31], may also apply in the future, and may even supplant the relative theory.

→ *Back to 2.8 Open Research Data*

[27] ATF 136 II 508, c. 3.2.

[28] Avis 05/2014 sur les techniques d'anonymisation, adopté le 10 avril 2014 par l'ancien groupe de travail «Article 29» sur la protection des données

[29] CJUE, C-582/14, Patrick Breyer contre Bundesrepublik Deutschland du 12 mai 2016, par. 46.

[30] BAERISWYL, Bruno, Big Data zwischen Anonymieserung und Re-Individualiserung, in: Big Data und Datenschutz – Gegenseitige Herausforderungen, Schulthess, 2014, p. 54; CELLINA Eva, La commercialisation des données personnelles, Schulthess, 2020, § 150; CICHOKI Michal, Big Data und Datenschutz: Ausgewählte Aspekte, in: Jusletter IT 21 mai 2015, N 19; GASSER Urs, Perspectives on the Future of Digital Privacy, in: Revue de droit suisse (RDS), Congrès de la société suisse des juristes, 2015, p. 351; MEIER Philippe, Le défi de Big Data dans les relations entre privés: Avec quelques réflexions de lege ferenda, in: Big Data et droit de la protection des données, Schulthess, 2016, p. 54; WEBER Rolf H., Herausforderungen für das Datenschutzrecht, in : Big Data et droit de la protection des données, Schulthess, 2016, p. 10; WEBER Rolf H., Big Data: Rechtliche Perspektive, in: Big Data und Datenschutz – Gegenseitige Herausforderungen, Schulthess, 2014, p. 20.

[31] Rosenthal David/Jöhri Yvonne (éd), Handkommentar zum Datenschutzgesetz, Schulthess, 2008, Rosenthal ad art. 3 LPD N 24 ; Meier Philippe, Protection des données, Fondements, principes généraux de droit privé, Stämpfli, 2011, N 445 ; Meier Philippe, Le défi de Big Data dans les relations entre privés: Avec quelques réflexions de lege ferenda, in: Big Data et droit de la protection des données, Schulthess, 2016, p. 55; Avis 05/2014 sur les techniques d'anonymisation, adopté le 10 avril 2014 par l'ancien groupe de travail «Article 29» sur la protection des données. Furthermore, in an obiter dictum, the Federal Administrative Court supported the absolute theory (ATAF A-6/2015 of July 26, 2017, c. 6.4.3.2): «[...] [W]hile it would theoretically be possible to break the established code - and thus indirectly identify - by means of another database[,] the risk exists that the person's identity can be re-established without disproportionate efforts.»