Research Data Protection

Good practices for sensitive & personal data

Audrey.Bellier@unige.ch (Arve)
Floriane.Muller@unige.ch (CMU)

December 2018
AGENDA

- Specificity of some kind of data
- What it means: regulations, responsibilities
- Good practices (Ethical, IT, IP)
  - Consent and data subject rights
  - Anonymization
  - IT protection & secure transmission
CLASSIFICATION OF RESEARCH DATA

observational
derived / compiled
simulation / models
experimental
reference
NOT ALL DATA IS EQUAL

• **Personal data:** “all information relating to an identified or identifiable person;”
  
  (Swiss Federal Act on Data Protection)

• **Sensitive personal data:** “data on:
  1. religious, ideological, political or trade union-related views or activities,
  2. health, the intimate sphere or the racial origin,
  3. social security measures,
  4. administrative or criminal proceedings and sanctions;”
  
  (Federal Act on Data Protection)

• **Confidential / «critical» data**
Personal, sensitive or critical data?

https://www.socrative.com/
(SENSITIVE) PERSONAL DATA

Image reused from CESSDA ERIC Expert Tour Guide on Data Management which is licensed under a CC-BY-SA 4.0 International License
SPECIFIC DATA -> SPECIFIC REGULATIONS

LPD

LRH

LIPAD

Funder requirements

GDPR

Ethics committee

Please Stay on the Path

Photo CC0 by Mark Duffel on Unsplash
By collecting and/or handling personal data, researchers are thereby required to follow a number of principles including:

- **Transparency** - processing personal data “lawfully, fairly and in a transparent manner”
- **Data Minimization** - data use shall be limited to the purpose of the respective research
- **Accuracy** - inaccurate data must be “erased or rectified without delay”
- **Integrity and Confidentiality** - data must be protected by appropriate security measures (technical and organizational)
**Federal Act on Data Protection (FADP)** [Loi fédérale sur la Protection des données - LPD]
- Defines what is considered (sensitive) personal data.
- Defines general data protection provisions
  - Such as: proportionality, correctness of data, possibility of cross-border disclosure, of data processing by third parties,…
- Is complemented by **Ordinance to the Federal Act on Data Protection** OLPD

**Loi sur l’information du public, l’accès aux documents et la protection des données personnelles** (LIPAD)
- Processing for general purposes (art. 41, LIPAD)

*Public institutions are entitled to process personal data for general statistical purposes, scientific research (…) under the conditions that:*

- (b) such data are **destroyed or made anonymous** as soon as the purpose of the specific processing operation concerned so permits;
- (c) the data collected for these purposes alone are **not communicated to any other institution**, entity or person;

(LIPAD, art. 41, our translation)
FEDERAL LAW ON HUMAN RESEARCH (LRH)

Applies to research on human diseases and on the structure and functioning of the human body, carried out:
- on people;
- on deceased people;
- on embryos and fetuses;
- on biological material;
- on personal data related to health.

Does not apply to the research carried out:
- on *in vitro* embryos (within the meaning of the Federal Act of 19 December 2003 on embryonic stem cell research)
- on anonymous biological material;
- on health-related data that have been collected *anonymously or anonymized*.

Cantonal Ethics Commission (CCER)

http://ge.ch/sante/commission-cantonale-dethique-de-recherche-ccer/commission-cantonale-dethique-de-recherche-ccer
Research projects involving human participants conducted at the University that are not under the responsibility of the Cantonal Ethics Commission (CCER)

Commission universitaire d’éthique (CUREG)
https://www.unige.ch/commissionethique/

Faculty-specific ethics committees:

• **Psychologie et Sciences de l’éducation**
  https://www.unige.ch/fapse/faculte/organisation/commissions/commission-ethique/

• **Traduction et Interprétation**
  https://www.unige.ch/fti/fr/faculte/organisation/commissions/ethique/

• **Sciences de la Société**
  https://www.unige.ch/sciences-societe/faculte/organisation/commissions-de-la-faculte/commission-dethique/
2.1 How will ethical issues be addressed and handled?
- What is the relevant protection standard for your data? Are you bound by a confidentiality agreement?
- Do you have the necessary permission to obtain, process, preserve and share the data? Have the people whose data you are using been informed or did they give their consent?
- What methods will you use to ensure the protection of personal or other sensitive data?

2.2 How will data access and security be managed?
- What are the main concerns regarding data security, what are the levels of risk and what measures are in place to handle security risks?
- How will you regulate data access rights/permissions to ensure the security of the data?
- How will personal or other sensitive data be handled to ensure safe data storage and transfer?

2.3 How will you handle copyright and Intellectual Property Rights issues?
- Who will be the owner of the data?
- Which licenses will be applied to the data?
- What restrictions apply to the reuse of third-party data?

4.2 Data Sharing and reuse: Are there any necessary limitations to protect sensitive data?
SOME GOOD PRACTICES

Apply for processing AND sharing

• Upheld your ethical responsibilities and protect your data subject rights

• Anonymize when possible

• Use appropriate IT Protection measures
**CONSENT: LEGAL REGULATIONS (LRH)**

Art. 7 Consent

1. Research on human beings may only be carried out if the person concerned has given **informed consent** or if he or she has not exercised his or her right of objection after having been informed in accordance with this Law.

2. The person concerned **may at any time refuse to participate** in a research project or **revoke his or her consent** without having to justify his or her decision.

Art. 16 Informed consent

2. The following information must be provided to the data subject orally and in writing, in an understandable form:
   a. the nature, purpose, duration and conduct of the research project;
   b. the foreseeable risks and constraints;
   c. the expected benefit of the research project, in particular for itself or others;
   d. the measures to ensure the protection of his or her personal data;
   e. its rights
CONSENT FORMS: GOOD PRACTICES

✓ Obtain the informed consent of participants **before** involving them in a research project.

✓ Give them the conditions to make an **informed decision**
  - Enough **information** (cf. LRH, art. 16)
  - Presented in a **clear** and appropriate manner
  - Enough **time** to think

✓ Anticipate the impact of a **possible retraction** and clarify them:
  - What it implies for the data collected until then?

Examples

http://www.who.int/rpc/research_ethics/informed_consent/en/
https://www.swissethics.ch/templates_e.html
CONSENT FOR DATA ARCHIVING AND SHARING

Identify and explain the possible future uses of the data and offer the participant the option to consent on a granular level.

For example, in a qualitative study:

I agree to:

- the non-anonymised audio recording of my interview being archived and disseminated for reuse
- the anonymised transcript of my interview being archived and disseminated for reuse
- any photographs of me taken during interview being archived and disseminated for reuse

ANONYMISATION

- **Avoid** collecting any identifiable information at all! (or minimize → reduce the time and effort required to anonymise it later)

- Personal data: **direct identifiers** (name, adress, phone #,...) vs. **indirect identifiers** (when they are placed with other information could also reveal an individual)

- **Anonymisation:** irreversibly destroys any way of identifying the data subject (does not fall under regulations such as GDPR, LRH....)

- **Pseudonymisation:** allows to re-identify the data subject with additional information (still under regulations)

This may involve **removing or aggregating variables**

Aggregate or **reduce the precision** of a variable such as age or place of residence

**Restrict the upper or lower ranges** of a continuous variable to hide outliers if the values for certain individuals are unusual or atypical within the wider group researched

---

ANONYMISATION OF QUALITATIVE DATA

- **Plan anonymisation** at the time of transcription or initial write-up

- Using pseudonyms or generic descriptors, rather than blanking-out that information

- Use pseudonyms that are consistent within the research team and throughout the project

- Use 'search and replace' techniques carefully so that unintended changes are not made, and misspelt words are not missed

- Identify replacements in text clearly:
  
  **[brackets]** or using XML tags such as `<seg>`word to be anonymised`</seg>`

- Create an **anonymisation log** of all replacements, aggregations or removals made and store such a log **securely and separately** from the anonymised data files
ANONYMISATION TOOLS

Amnesia is a data anonymization tool, that allows to remove identifying information from data.

This text anonymisation helper tool can help you find disclosive information to remove or pseudonymise in qualitative data files.
Data Security Basics

Evaluate your security needs for
- Physical data & digital data
  - eg., patient records, signed consent forms, interview cover sheets…
- Depending on the sensitive nature or not of the data

Ordinance to the Federal Act on Data Protection, art 8-10 Technical and organisational measures, states that one has to
“ensure the confidentiality, availability and the integrity of the data in order to ensure an appropriate level of data protection. In particular, […] protect the systems against the following risks:
  a. unauthorised or accidental destruction;
  b. accidental loss;
  c. technical faults;
  d. forgery, theft or unlawful use;
  e. unauthorised alteration, copying, access or other unauthorised processing.”
DATA ACCESS CONTROLS

In addition to regular security measures such as back-up, etc, Consider:

• Controll access to rooms and buildings
• Lock computers systems with passwords
• Do not store confidential data on servers or computers connected to an external network
• Log the access to hard-copies / digital copies
• Implement control access to data files: e.g. no access, read only, read and write, administrator-only permission
• Encrypt sensitive data before sharing it with authorized people.
• Transport sensitive data only under exceptional circumstances, even for repair purposes
• Destroy data in a consistent manner when needed!
**Encryption**

- process that encodes a message or file so that it can be only be read by certain people.
- uses an algorithm to scramble data and then uses a key for the receiving party to unscramble the information.

![Symmetric Encryption](https://commons.wikimedia.org/wiki/File:Symmetric_encryption.png)

- Use Reliable encryption software:
  Eg. GnuPG (Gnu Privacy Guard) [recommended by the UK data archive] or 7-zip

https://www.ukdataservice.ac.uk/manage-data/store/encryption
TAKE HOME MESSAGE

 ✓ Discuss whether you actually need to collect personal data to carry out your research
 ✓ Consider collecting data anonymously if possible
 ✓ Identify which personal data will be included in your research
 ✓ Include aspects of data protection in your DMP
 ✓ Create and use consent forms
 ✓ Find the appropriate anonymisation strategy for your research
 ✓ Securely store, control access, and transfer your data
Specific Questions / Cases?

https://www.unige.ch/researchdata

UNIGE at your Service: MENU

Trainings & Lectures
DMP Reviewing
Contact points
UNIGE policy
Help & Support

Help & Support

You want your DMP proofread before submitting it to the funding agency? Send it to us for review via our ticketing tool below and we will give you our comments as soon as possible.

You can of course send us any questions you may have about research data. We are committed to responding quickly.

Ask a question or Submit your DMP for comment
THANK YOU FOR YOUR ATTENTION

http://www.unige.ch/biblio

Audrey.Bellier@unige.ch
Floriane.Muller@unige.ch