Want to collect personal data?

An introduction to processing personal data for research purposes

Eva J. B. Payne, MPhil, Adviser, NSD
Ina Nepstad, PhD, Senior Adviser, NSD

23 September 2020

cessda.eu @CESSDA_Data NORWEGIAN CENTRE FOR RESEARCH DATA

Doi: 10.5281/zenodo.4046051

Presentation

- What is processing? An overview of the data life-cycle
- What is personal data? An overview of important terms
- What is the GDPR and why is it relevant for you as a researcher?
- What is a legal basis for processing?
- Which rights apply and what do they entail?
- What are the principles for processing personal data?
- Social benefit of research in relation to potential risk for your participants

- Useful tips along the way!

What is “processing”?

Processing entails any operation which is performed on personal data, such as…

- collecting,
- registering,
- downloading,
- recording,
- structuring,
- combining,
- storing,
- sharing,
- transferring,
- publishing,
- anonymizing,
- deleting etc.

i.e. anything you can “do” with data

Data life-cycle

Data management:

- how data will be collected

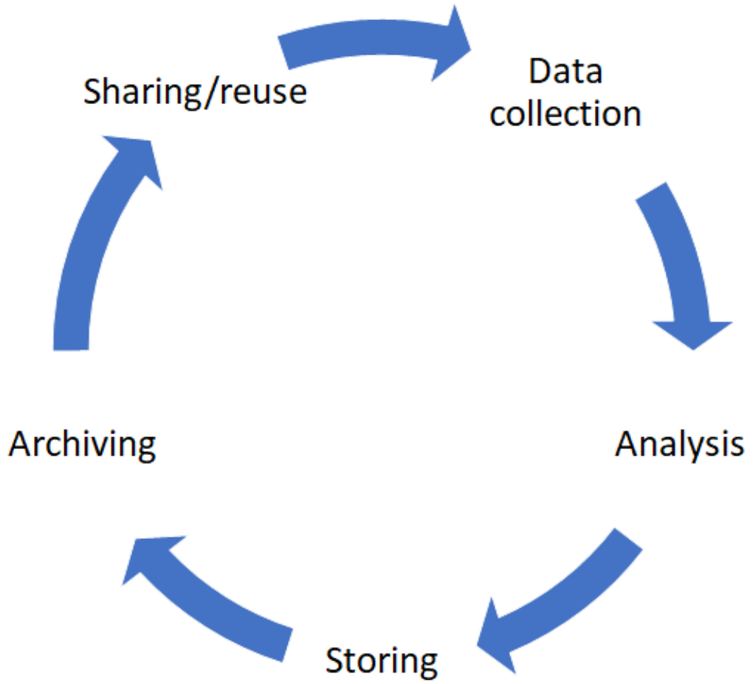
- how collected data will be organised/structured/analysed

- how data will be stored

- whether the data will be shared with others and who will have access (during research project and afterwards)

- whether data will be archived

- how data will be used in the future



TIP 1:

Plan ahead

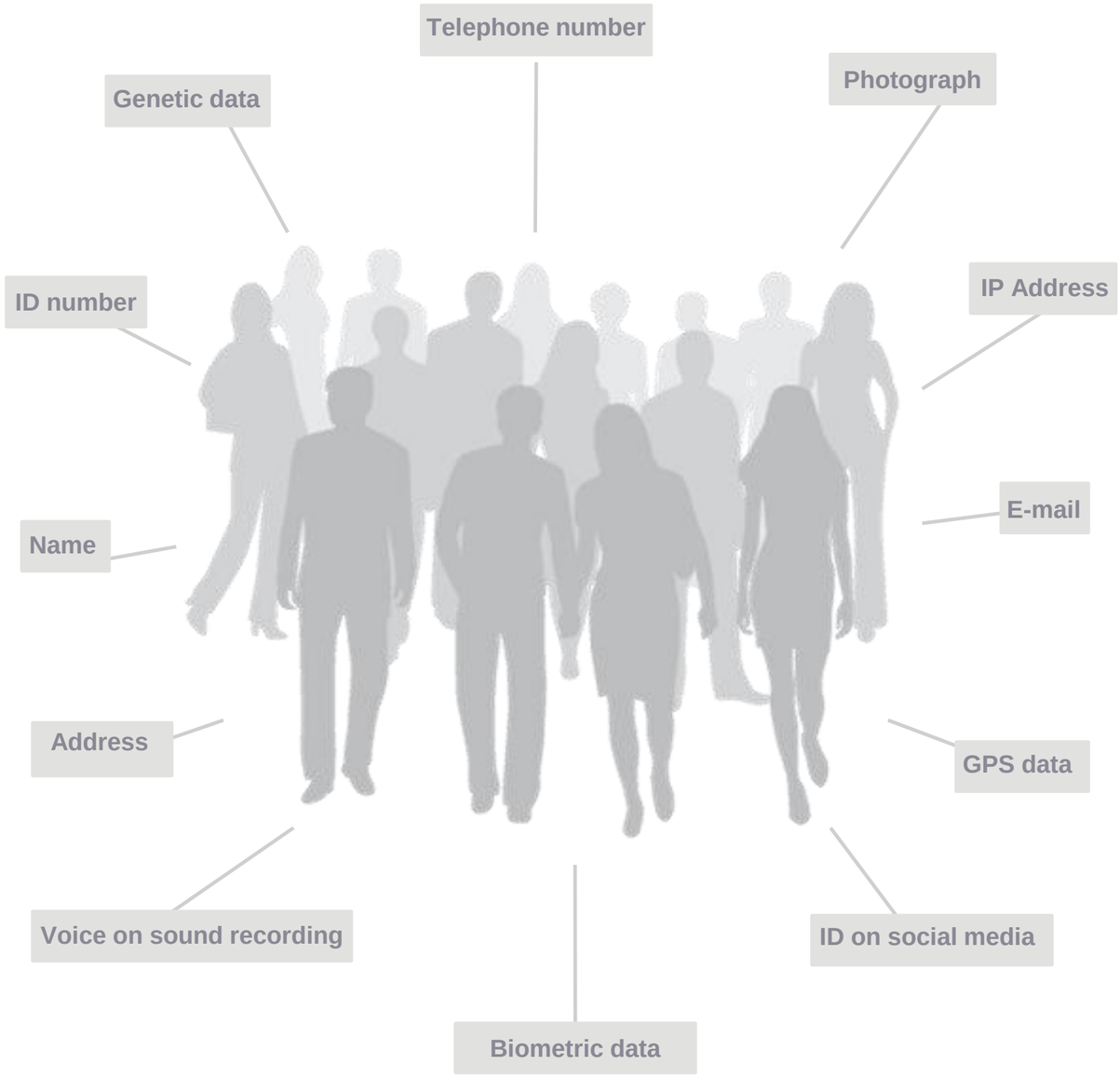
During the **planning stages** think through the life-cycle of your collected data

Be prepared to collect good research data!

What is personal data?

- Directly identifiable
- Indirectly identifiable

Personal data means any information relating to an identified or identifiable person



What is personal data?

Special categories of personal data

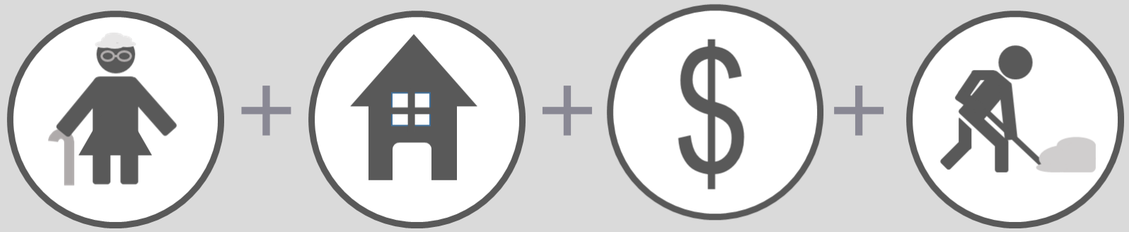
- Racial or ethnic origin
- Political opinions
- Religious beliefs
- Philosophical beliefs
- Trade Union Membership
- Health data
- Sex life or sexual orientation
- Criminal convictions and offences

What is personal data?

- Directly identifiable
- Indirectly identifiable

A person can be identified based on a combination of background information/demographic data

e.g. gender, age, workplace, income, nationality etc.

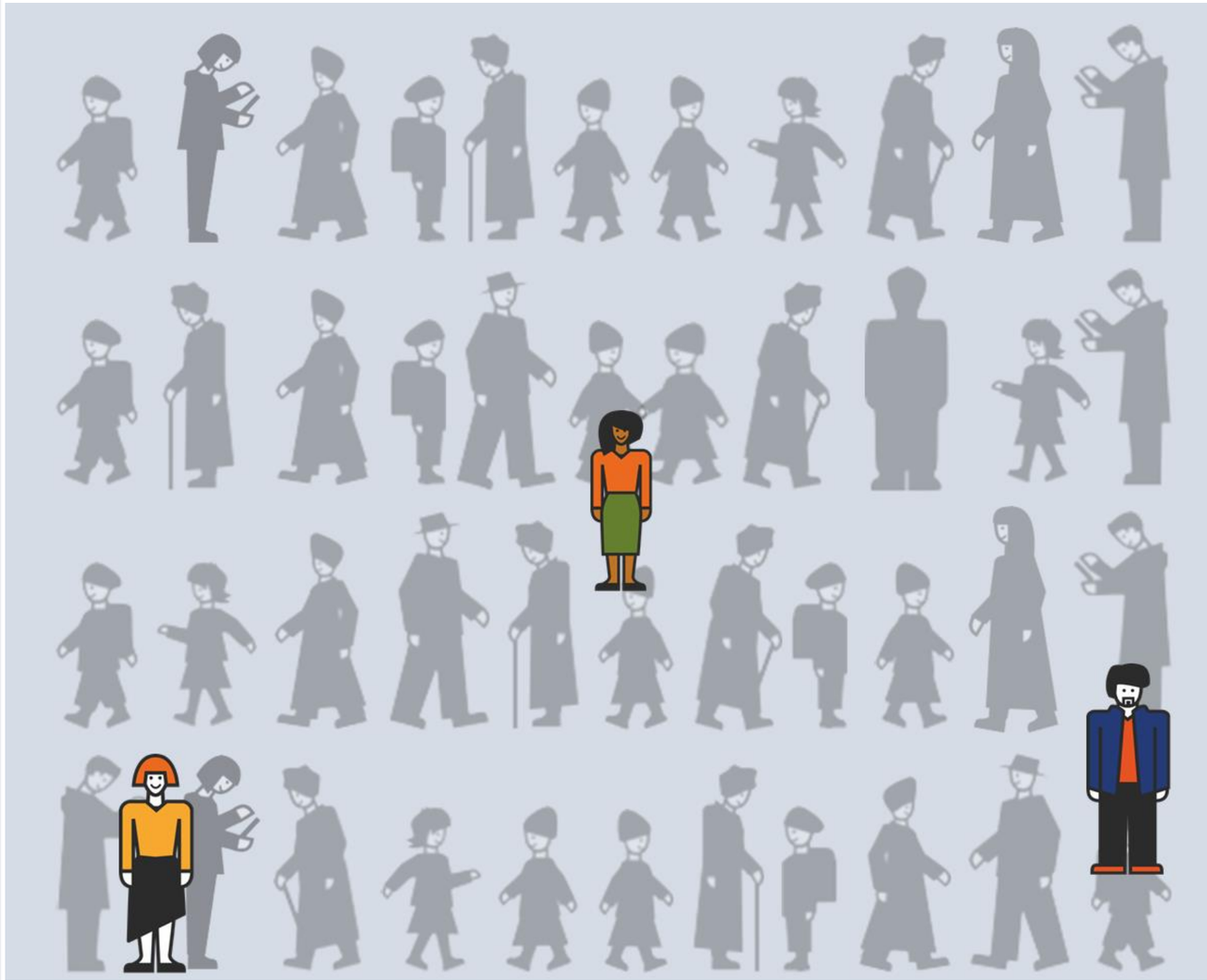


**Topic and context,
sample size and criteria,
and type of background data**

What is personal data?

- Directly identifiable
- Indirectly identifiable

Diagnosis

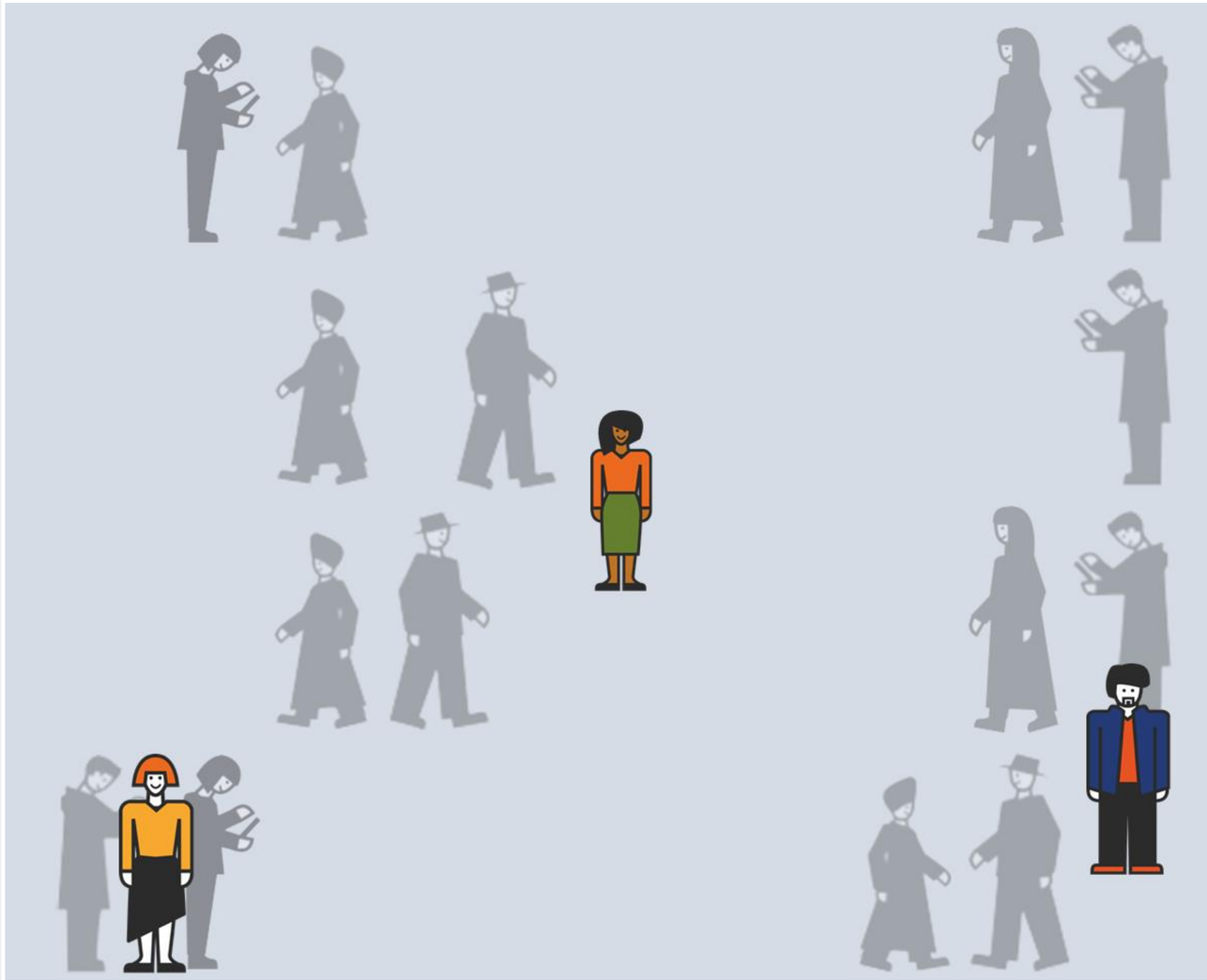


What is personal data?

- Directly identifiable
- Indirectly identifiable

Diagnosis

Age



What is personal data?

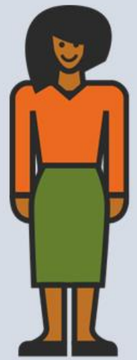
- Directly identifiable
- Indirectly identifiable

Diagnosis

Age

Occupation and place of work

Income



What is anonymous data?

Information that can in no way be linked to an individual person

- Directly
- Indirectly
- through a list of names/codes (i.e. scrambling key)

What is anonymous data?

Even if the data you are analyzing does not identify individual persons….

….if a list of names/codes exists that makes it possible to identify those person then your data is

personal data

It is called

pseudonymised (personal) data

What is the GDPR and why is it relevant for you as a researcher?

harmonises data protection legislation in Europe and gives control to individuals over their personal data

the protection of natural persons in relation to the processing of personal data is a fundamental right

enables the free flow of personal data between Member States, whilst also ensuring a high level of personal data protection

TIP 2:

Think «legal basis, rights and principles»

If you consider these three aspects of processing personal data

- before data collection

- and before each stage of your project

then you will have covered what is most important from a data protection perspective

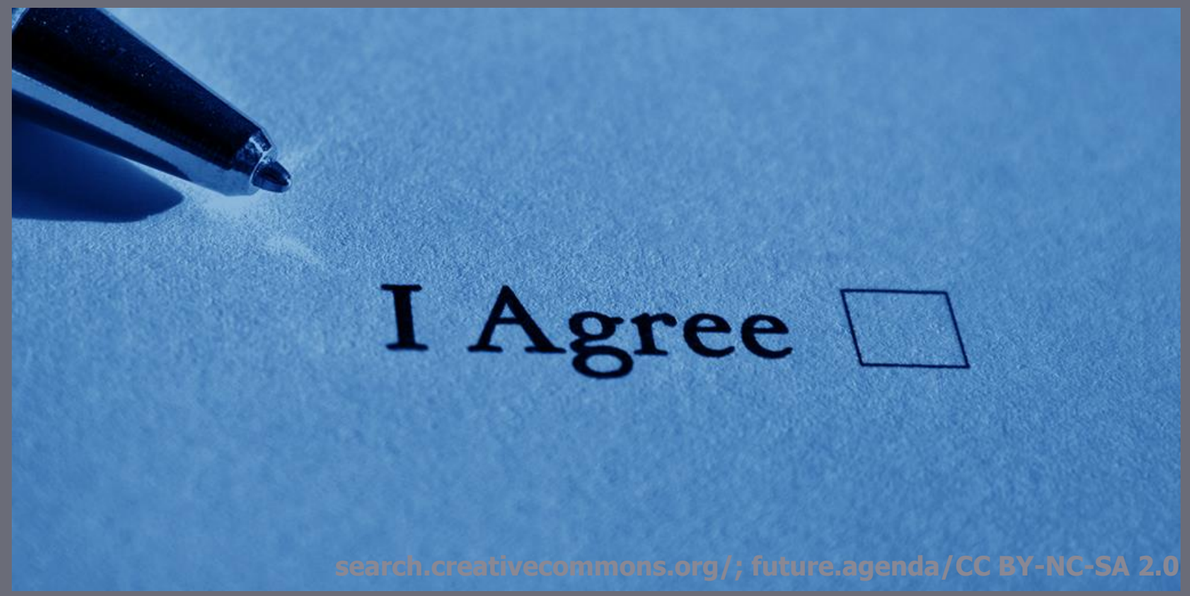
What is a “legal basis” for processing?

Legal grounds for processing personal data

Processing is lawful only if certain conditions/grounds apply

Legal bases are found in (GDPR):
- Article 6 (general categories)
- Article 9 (special categories)

Consent / Explicit Consent



Article 6(1)(a) / Article 9(2)(a)

For consent to be valid it must be:

- Freely given
- Specific
- Informed
- Unambiguous statement or action
- Documented, i.e. can be demonstrated
- As easy to withdraw as to give

See articles 4 (11) and 7 in GDPR

Public interest /

Archiving purposes in the public interest

Scientific or historical research purposes or statistical purposes

Article 6(1)(e) / Article 9(2)(j)

What rights do data subjects have?

What do these rights entail?

Information ensures fair and transparent processing

Information provided should meet requirements for:

Form and **Content**

See articles 12, 13 and 14

The data subject's **right to be informed**

Your **obligation** to provide information

Content

- ✓ which institution is responsible for the project (the data controller)
- ✓ contact details for institution (project leader) and the data protection officer (if applicable)
- ✓ the purposes of processing personal data and legal basis for processing
- ✓ who will have access to/receive the personal data (e.g. project group, external researchers, data processors)
- ✓ if applicable, that personal data will be transferred to a third country or international organisation, and the legal basis for transfer (including which safeguards will protect the data)
- ✓ the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period
- ✓ what rights the data subjects have and how they exercise their rights
- ✓ if processing is based on consent: the right to withdraw consent at any time

Other rights….

See articles 15-21 in GDPR

right of access

right to rectification / correction of incorrect personal data

right to erasure / deletion

right to restrict processing

right to data portability (a copy)

right to object to processing

rights in relation to automated decision making and profiling

right to lodge a complaint with the supervisory authority

Rights apply so long as the data subject can be identified in the collected data

Exemptions from rights must be justified and must have a legal basis

TIP 3:

Be realistic.

Don’t limit yourself unnecessarily.

Based on our experience researchers often underestimate how long they will need to achieve their research purposes.

Researchers find data protection legislation challenging **BUT** it is not necessary to delete all your collected data at the end of the project.

Anonymised data can (and often should) be archived for future research purposes.

Personal data can also be archived – **think «legal basis, rights and principles»**

There should be a **reason** why personal data (and not anonymous data) is being archived, and the **type of data** must be taken into account

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitations
- Integrity and confidentiality
- Accountability

7 principles of the GDPR

Lawfulness, fairness and transparency

–

Processing of personal data must happen in a *lawful way* and thus have a legal basis which makes the processing legitimate

Fairness means that your actions must match up with how it was described to data subject

A clear notice is what the concept of *transparency* is about

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitations
- Integrity and confidentiality
- Accountability

7 principles of the GDPR

Purpose limitation

—

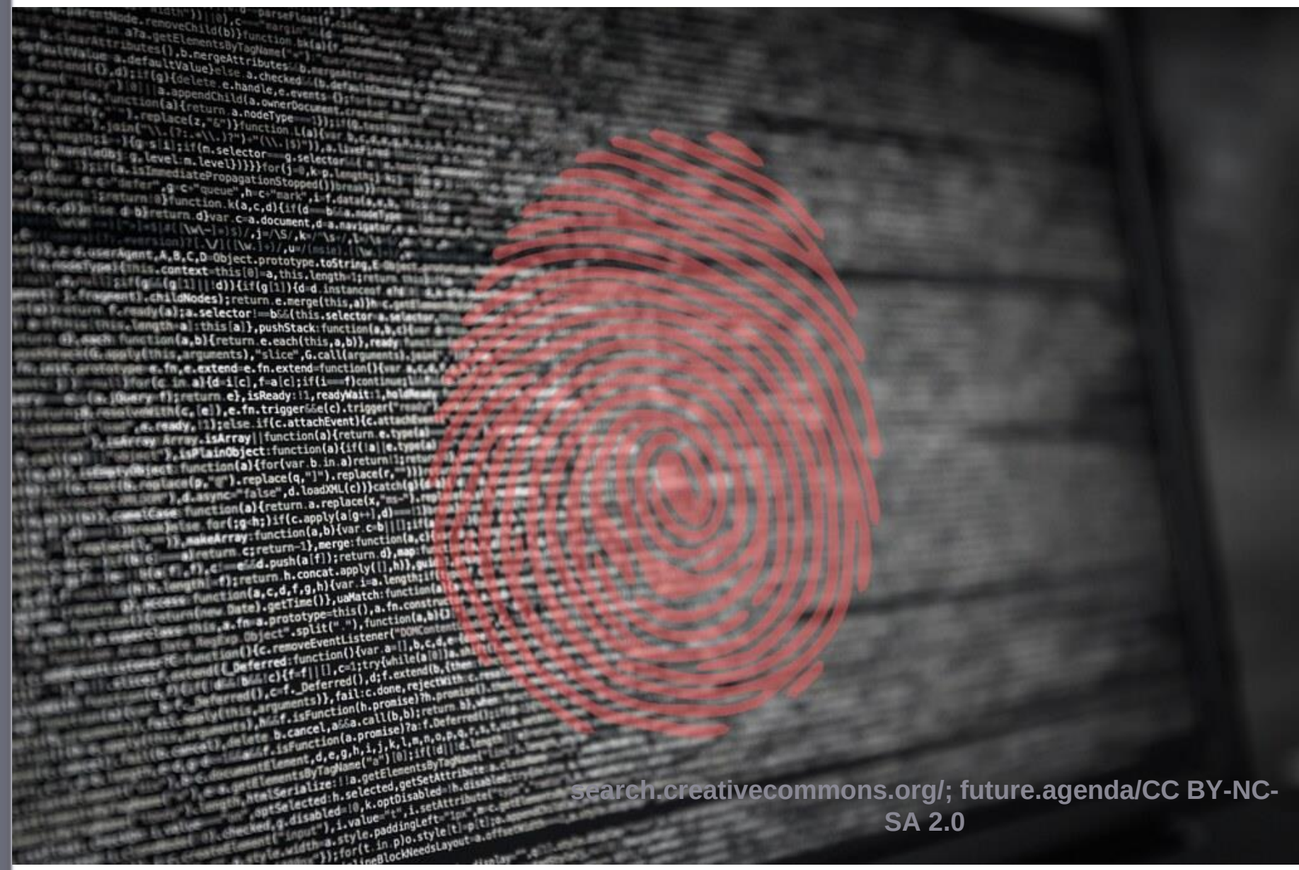
Be specific

You must inform the participants about the purpose of the data collection

“specified, explicit and legitimate”

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitations
- Integrity and confidentiality
- Accountability



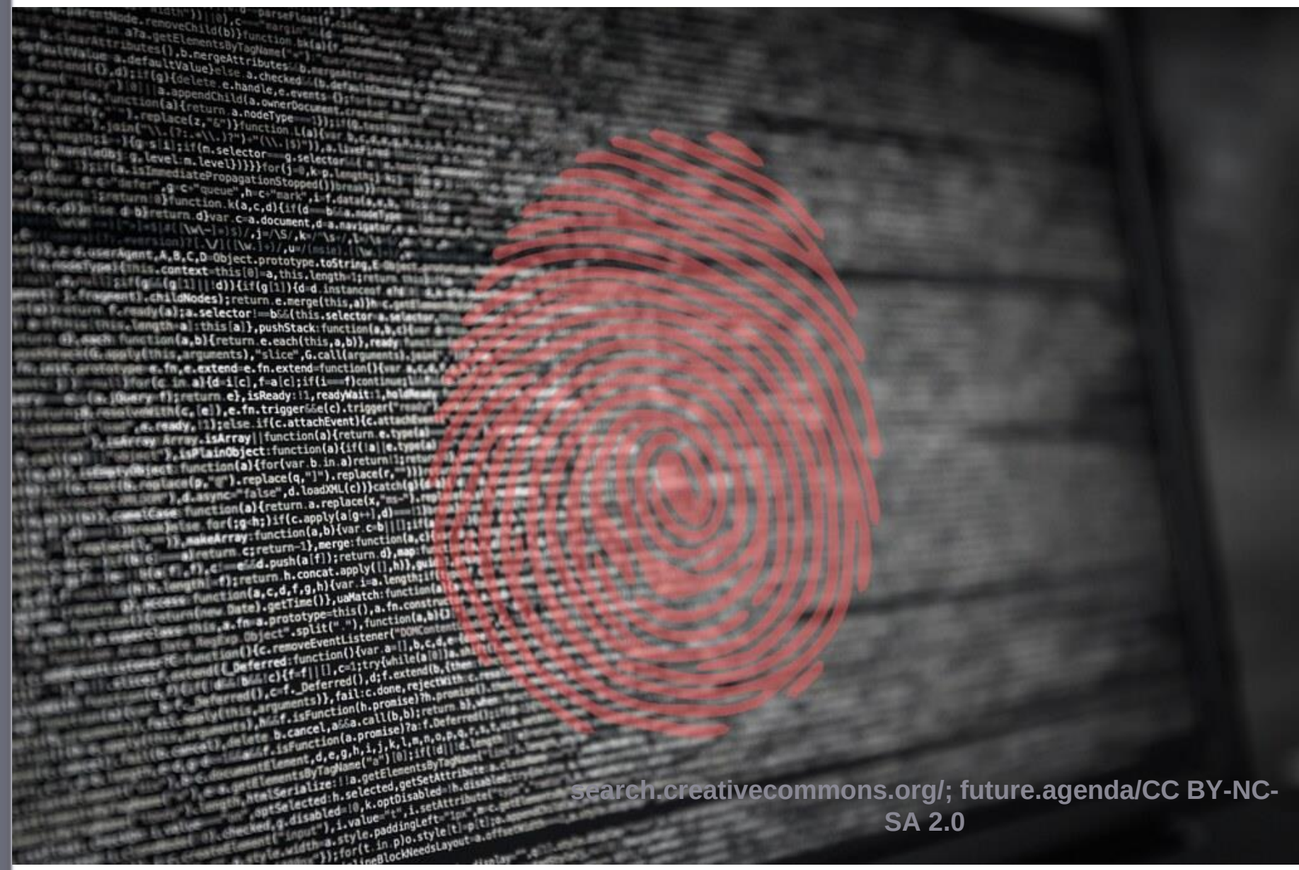
7 principles of the GDPR

Data minimization

—

Collect the minimum data you need

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



7 principles of the GDPR

Lawfulness, fairness and transparency

Purpose limitation

Data minimization

Accuracy

Storage limitations

Integrity and confidentiality

Accountability

7 principles of the GDPR

Accuracy

—

Your data must be accurate and up-to-date

"accurate and where necessary kept up to date"

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- **Storage limitations**
- Integrity and confidentiality
- Accountability

7 principles of the GDPR

Storage limitations

—

Retain the data for the necessary period and then erase or anonymize

"kept in a form which permits identification of data subjects for no longer than necessary"

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitations
- Integrity and confidentiality
- Accountability

7 principles of the GDPR

Integrity and confidentiality

—

Keep the data secure!

"in a manner [ensuring] appropriate security", which include "protection against unlawful processing or accidental loss, destruction or damage".

7 principles of the GDPR

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitations
- Integrity and confidentiality
- Accountability

7 principles of the GDPR

Accountability

—

You are responsible for compliance with the principles of the GDPR

The social benefit vs. risk/disadvantage for data subjects

Risk to the rights and freedoms of data subjects depends on, i.a.:

- how sensitive the data is
- how easy it is to identify individuals
- the quantity of personal data
- how securely the data is being stored

Ways of reducing risk are therefore limiting the amount of sensitive data and making individuals less identifiable

TIP 4:

Be organized and have a system

When collecting, storing and analysing data don’t take anything for granted….

- expect to forget which interviewee is which
- store names (directly identifiable data) separately from other data
- keep your metadata

Think about **FAIR** principles in advance. That your future data should be:

- Findable
- Accessible
- Interoperable
- Reusable

Thank you for listening!

Eva J. B. Payne, MPhil, Adviser, NSD
Ina Nepstad, PhD, Senior Adviser, NSD
NSD – Norwegian Centre for Research Data AS

cessda.eu

@CESSDA_Data

Questions & answers

cessda.eu

@CESSDA_Data