

# Protect

---

Data Management Expert Guide (DMEG):

*Anne-Mette Somby, NSD*

*CESSDA Train the Trainers Workshop  
14.May 2019 Athens*

 [cessda.eu](https://cessda.eu)  [@CESSDA\\_Data](https://twitter.com/CESSDA_Data)

# Overview

◊ *This part of the tour guide focuses on key legal and ethical considerations in creating shareable data*

- ◊ Ethical review
- ◊ Data Protection
- ◊ Informed consent
- ◊ Anonymization

# Disclaimer

- ◊ The information in this presentation is based on our current interpretation of the legislation and its implications for research and the archiving of research data
- ◊ This is a very fluid area and thus changes are still possible

# Ethical review

# Ethical Review Process

- ◊ Ethical review is about helping you as a researcher to think through the ethical issues surrounding your research
- ◊ The principles of good research practice encourage you to consider the wider consequences of your research and engage with the interest of your participants
- ◊ Ethics review by a Research Ethics Committee (REC) is typically required when (sensitive) personal data are being collected
- ◊ Among other duties, this involves ensuring that research complies with national and international data protection laws regarding the use of personal information collected in research

# Ethical Arguments for Archiving Data

- ◊ Not burden over-researched, vulnerable group
- ◊ Make best use of hard-to-obtain data, e.g. elites, socially excluded, over- researched
- ◊ Extend voices of participants
- ◊ Provide greater research transparency
- ◊ *In each, ethical duties to participants, peers and public may be present*

# Data Protection

# The General Data Protection Regulation (GDPR)

- ◊ The GDPR applies from the 25 May 2018
- ◊ The GDPR applies to any data controller or data processor in the EU who collects personal data about a data subject of any country, anywhere in the world
- ◊ A data controller or data processor that is based outside the EU but collects personal data on EU citizens will also be covered by the GDPR
- ◊ This means that a researcher (data controller) based within the EU who collects personal data about a participant, from any other country within the EU, or the world, needs to comply with the GDPR
- ◊ Also means a researcher (data controller) outside the EU who collects personal data about a participant in the EU will be covered when this relates to offering goods/services or the monitoring of their behaviour within the EU



# GDPR applies to

- ◊ The GDPR applies only to 'personal data' and data of 'living persons'
- ◊ Data which do not count as personal data do not fall under data protection legislation
- ◊ Though there may still be ethical reasons for wanting to protect this information!
- ◊ Two categories of personal data , general and special categories

# What is personal data?



# Short definition “personal data”

◇ ***Personal data is any information that may be used to identify a person directly or indirectly***

◇ **Directly identifying personal data**

*through full name, personal identification number*

◇ **Indirectly identifying personal data**

*through a combination of background information*

# List of special categories

- ◇ Racial or ethnic origin
- ◇ Religious opinion, political and philosophical beliefs
- ◇ Health
- ◇ Trade union membership
- ◇ Sexual orientation or sex life

# Legal grounds for processing

- ◊ All processing of personal data requires legal basis. The most common for research are:
- ◊ (Article 6): a) consent , e) necessary for the performance of a task carried out in the public interest
- ◊ Special categories of data (Article 9):
- ◊ Prohibited unless: a) explicit consent e) personal data are manifestly made public by the data subject j) necessary for archiving, scientific or statistical purposes

# Other legal grounds

- ◇ Performance of a contract (art. 6 nr. 1 c)
- ◇ Purposes of legitimate interests (art. 6 nr. 1 f)
- ◇ Data which are manifestly made public by the data subject (art. 9 nr. 2 e)
- ◇ Purposes of preventative or occupational medicine (art. 9 nr. 2 h)

# Principles Relating to Processing of Personal Data

## 1. Process lawfully, fair and transparent

The participant is informed of what will be done with the data and data processing should be done accordingly

## 2. Keep to the original purpose

Data should be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. \*Note Article 5(1)(b), Article 89 \*

## 3. Minimise data size

Personal data that are collected should be adequate, relevant and limited to what is necessary

# Principles Relating to Processing of Personal Data

## 4. Uphold accuracy

Personal data should be accurate and, where necessary kept up to date. Every reasonable step must be taken to ensure that personal data that are inaccurate are erased or rectified without delay

## 5. Remove data that are not used

Personal data should be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed. \* Note Article 5(1)(e), Article 89 \*

## 6. Ensure data integrity and confidentiality

Personal data are 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



# GDPR Archiving and Research Exemption

- ◊ *Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that **technical and organisational measures** are in place in particular in order to ensure respect for the principle of **data minimisation**. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner*

# GDPR Archiving and Research Exemption

- ◊ Principle 2 (purpose) and 5 (anonymization) are less strict

This means that further processing is allowed

- And that personal data can be stored for longer periods
- if necessary safeguards are in place

# The Grounds for Processing Personal Data

- ◆ There are 6 grounds for the processing of personal data, and one of these must be present in order to process a data subject's personal data:

1. Consent of the data subject
2. Necessary for the performance of a contract
3. Legal obligation placed upon controller
4. Necessary to protect the vital interests of the data subject
5. Carried out in the public interest or is in the exercise of official authority
6. Legitimate interest pursued by controller

# GDPR- Data Subject Rights

- ◊ transparency (art. 12), **information (art. 13), access (art. 15)**, rectification (art. 16), **erasure (art. 17)**, restriction of processing (art. 18), notification (art. 19), data portability (art. 20).

# Best Practise for Legal Compliance

- ◊ Investigate early which laws apply to your project
- ◊ Do not collect personal and sensitive data if not relevant to your project
- ◊ Seek advice from your research office
- ◊ If you need personal and sensitive data:

Inform and get consent from the participants

You may need an ethical and legal assesment

# Strategy for Sharing Data

- ◊ Obtain informed consent, also for data sharing and preservation or curation
- ◊ Protect identities e.g. anonymisation and not collecting personal data if not necessary
- ◊ Regulate access where needed (all or part of data) e.g. by group, use or time period
- ◊ Securely store and protect personal and sensitive data

# DPIA - Data Protection Impact Assessment

- Sensitive data
- Consent not possible
- Long term processing/ archiving
- Vulnerable group
- Very identifiable data
- ◊ Combination of the above
- ◊ The DPIA is a written document to be formally approved by the University and DPO

# Informed Consent



# Consent is needed Across the Data Lifecycle

- ◆ Participant Engagement in the research activities:

Interview, survey, observation

- ◆ Dissemination in presentations, publications and the web:

Consent for use of quotes for articles and video publicity

- ◆ Data sharing and archiving

# Informed Consent

- ◊ Consent needs to be **freely given, informed, unambiguous, specific** and by a **clear affirmative action** that signifies agreement to the processing of personal data
- ◊ When special categories data are processed – and the processing grounds for this is consent – there is a further requirement to the above that this must be based on **explicit consent**

# Informed Consent 2

- Freely given: must be a genuine choice, be able to refuse/withdraw without consequences, not be in a dependent relationship
- Specific - clear information on extent and consequences
- Informed: Content and form requirements, should be easily understood, easily accessible, clear and simple language, especially when the information is given to children
- Active: "opt in" - silence, pre-ticked boxes, and inactivity are not valid
- It should be as easy to withdraw consent as to give it
- Controller shall be able to demonstrate that consent is given (Article 7)

# Informed Consent- Data Sharing 1

- ◊ Adding the discussion of data sharing and archiving permits the participant to make an informed decision. This empowers them and puts them in charge of choosing whether they wish for their contribution to the research project – and their data – to be available for use in future research projects
- ◊ The best way to achieve informed consent for data sharing is to **identify** and **explain** the **possible future uses of their data** and offer the participant the option to consent on a **granular level**

# Informed Consent- Data Sharing 2

- ◊ For example, in a qualitative study, this may involve allowing the participant to consent to data sharing of the anonymised transcripts, the non-anonymised audio recordings and the photographs

# In Practice: Consent forms/Information Sheets

The interviews will be archived at ..... and disseminated so other researchers can reuse this information for research and learning purposes:

- I agree for the audio recording of my interview to be archived and disseminated for reuse
- I agree for the transcript of my interview to be archived and disseminated for reuse
- I agree for any photographs of me taken during interview to be archived and disseminated for reuse

# In Practice: Consent forms/Information Sheets

We expect to use your contributed information in various outputs, including a report and content for a website. Extracts of interviews and some photographs may both be used. We will get your permission before using a quote from you or a photograph of you.

After the project has ended, we intend to archive the interviews at .... Then the interview data can be disseminated for reuse by other researchers, for research and learning purposes.

# Timing and Form of Consent

	Advantage	Disadvantage
<b>Written consent</b>	<p>More solid legal ground, e.g. participant has agreed to disclose confidential info</p> <p>Often required by Ethics Committees</p> <p>Offers more protection for researcher (as they have written documentation of consent)</p>	<p>Not possible for some cases: infirm, illegal activities</p> <p>May scare people from participating (or have them think that they cannot withdraw their consent)</p>
<b>Verbal consent</b>	<p>Best if recorded</p>	<p>Can be difficult to make all issues clear verbally</p> <p>Possibly greater risks for researcher (in regards to adequately proving participant consent)</p>

	Advantage	Disadvantage
<b>One-off consent:</b> participant is asked to consent to taking part in the research project only once.	<p>Simple</p> <p>Least hassle to participants</p>	<p>Research outputs not known in advance</p> <p>Participants will not know all info they will contribute</p>
<b>Process consent :</b> participant's consent is requested continuously throughout the research project	<p>Ensures 'active' consent</p>	<p>May not get all consent needed before losing contact</p> <p>Repetitive, can annoy participants</p>



# Documenting Consent

- ◊ Under the GDPR, consent needs to be documented, which means (in the context of research) it will be important for researchers to maintain documented and accurate records of the consent obtained from their participants
- ◊ This could, for example, be written consent or audio recorded oral consent
- ◊ Though the GDPR does not require this consent to be in a written form, many UK research ethics committees and professional bodies do require this or recommend it as best practice

# Anonymisation

# Protection Participants

- ◊ The best way to protect your participant's privacy may be not to collect certain identifiable information at all
- ◊ The second best is anonymisation which allows data to be shared whilst protecting participant's personal information
- ◊ Anonymisation should be considered in the context of the whole project and how it can be utilised alongside, informed consent and access controls
- ◊ For example, if a participant consents to their data being shared then the use of anonymisation may not be required

# Identity Disclosure

◆ A persons identity can be disclosed through:

1. Direct identifiers e.g. name, address, postcode, telephone number, voice, picture g.
2. Indirect Identifiers

◆ Possible disclosure in combination with other information

e.g. occupation, geography, unique or exceptional values (outliers) or characteristics

# Anonymising Quantitative Data

- ◊ Remove direct identifiers e.g. names, address, institution, photo
- ◊ Reduce the precision / detail of a variable through aggregation / categorisation e.g. birth year instead of date of birth, occupational categories rather than job titles, area rather than village
- ◊ Generalise meaning of detailed text variable
- ◊ Restrict upper lower ranges of a variable to hide outliers
- ◊ Combining variables

# Anonymisation in Practice - Quant

Identifier type	Direct identifier	Strong indirect identifier	Indirect identifier	Anonymisation method
Full name	x			Remove/Change
Email address	x	x		Remove
Postal code			x	Remove/Categorise
Municipality of residence			x	Categorise
Municipality type			x	Remove
Video file displaying person(s)	x			Categorise
Year of birth		x		Categorise
Age			x	Categorise
Gender			x	
Marital status			x	
Occupation		(x)	x	Categorise
Employment status			x	
Ethnic group *		(x)	x	Categorise/Remove
Crime or punishment *			x	Categorise/Remove

# Anonymising Qualitative Data

- ◊ Plan or apply editing at time of transcription except: longitudinal studies - anonymise when data collection complete (linkages)
- ◊ Avoid blanking out; use pseudonyms or replacements
- ◊ Avoid over-anonymising – removing / aggregating information in text can distort data, make them unusable, unreliable or misleading
- ◊ Consistency within research team and throughout project
- ◊ Identify replacements, e.g. with [brackets]
- ◊ Keep an anonymisation log of all replacements, aggregations or removals made

# In Practice: Anonymisation - Quali

Ex 1. Health and Social Consequences of the Foot and Mouth Disease Epidemic in North Cumbria, 2001-2003 (study 5407 in UK Data Archive collection) by M. Mort, Lancaster University, Institute for Health Research.

Date of Interview: 21/02/02

Interview with **Lucas Roberts**, DEFRA field officer

Date of birth: **2 May** 1965

Gender: Male

Occupation: Frontline worker

Location: **Plumpton**, North Cumbria

**Lucas** was living at home with his parents, "but I'm hoping to move out soon" so we met at his parents' small neat house. We sat in a very comfortable sitting room with an open fire and **Lucas** made me coffee and offered shortbread. Although at first **Lucas** seemed a little nervous, quick to speech and very watchful he seemed to relax as we spoke and to forget about the tape.

Comment [v1]: Replace: Ken

Comment [v2]: delete

Comment [v3]: delete

Comment [v4]: Replace: Ken

Comment [v5]: Replace: Ken

Comment [v6]: Replace: Ken



# In Practice: Anonymisation - Quali

Yeah. So is part of your job to look for funding bids and to write funding bits or is that separate?

No. That was what P3 used to do and then it sort of passed down to... really it's with P1 and P4. But I don't actually think there's anything out there at the moment. I think at the moment, because there's all this money saving and things, there's nothing to... there isn't actually anything to access.

# What if Anonymisation is Impossible?

- ◊ *Anonymisation should be considered in the context of the whole project and how it can be utilised alongside, informed consent and access controls*

Obtain consent for sharing non-anonymised data

- » Regulate or restrict user access

# Group work

# Group work 1

- ◆ **Personal Data:**

- ◆ (10 minutes): Work in groups to propose solutions to the case study/studies.
- ◆ **Exercise 1:** A ph.d. will collect interview data that will contain personal information. She wants to make sure that she can share these data with new project partners in follow-up studies after her current ph.d. project. What are her options to make sure she can share her data?
- ◆ **Exercise 2:** A researcher's informants have not given their consent to the interviews being shared outside the researcher's current project. Can the researcher nevertheless archive the data (e.g. in a national archive for research data)?

# Group work 2

- ◆ Part 1 (5 minutes): Work individually to come up with at least one pro and one con for sharing your data. Then share and discuss the arguments you came up with in smaller groups.
- ◆ Part 2 (5 minutes): Share the different results with the other groups
- ◆ .



# Thank you for listening!

---

*Anne-mette.somby@nsd.no*

Contact us: [training@cessda.eu](mailto:training@cessda.eu)

Train the Trainers package:  
[www.cessda.eu/DMEG/TTT](http://www.cessda.eu/DMEG/TTT)

 [cessda.eu](http://cessda.eu)  [@CESSDA\\_Data](https://twitter.com/CESSDA_Data)



# More information?

[Anne-mette.somby@nsd.no](mailto:Anne-mette.somby@nsd.no)

<https://www.cessda.eu/Training>

