Handling Personal Data in Research

Handling Personal Data

Understanding the GDPR

- GDPR
- Definitions
- Lawfulness of processing : Informed consent
- Further Processing of research data
- Rights of a Data Subject

Applying the GDPR

- Sensitive personal Data (DPIA)
- Legal Documents
- Social Media
- Data Security
- Secure Storage / Encryption
- Pseudonymization & Anonymization
- Sharing Data

GDPR

General Data Protection Regulation

De Algemene verordening gegevensbescherming (AVG)

The 6 principles of the GDPR

The 6 principles of the GDPR

GDPR	Relevance to research
• Lawfulness	Obtaining informed consent
• Purpose limitation	Using personal data for research purposes
• Data minimization	The Rights of Data Subjects
• Data Accuracy	The Rights of Data Subjects
• Storage Limitation	Storing personal data for research purposes
• Integrity and Confidentiality	Securing personal Data

Definitions

Data Subject

- The individual whose data is being collected
- Natural Person

Personal Data

- Any information which can identify a natural person whether **directly or indirectly**

Sensitive Personal Data

- Personal information revealing a data subjects' ethnic origin, political opinions, religious beliefs, sexual orientation and genetic data or biometric data uniquely identifying a data subject

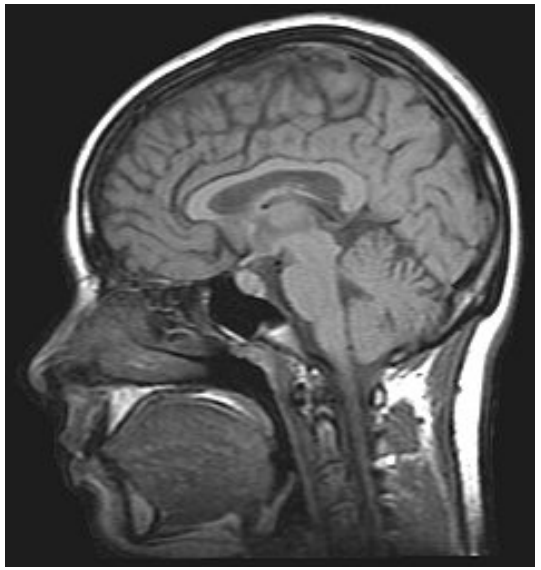
Personal Data

Is a First name, by itself, personal data?

Answer inTeams chat

Direct and Indirect identifiers

- Age
- First name
- IP address
- Last name
- Occupation
- Hair color
- Gender
- Instagram username
- Height
- Car color
- Political affiliation



Name	Occupation	Medical history	Religion	Political preference	Criminal history
xxxxxxxx	Prime minister of the Netherlands	Jaw-pains	Protestant	VVD	none
xxxxxxxx	Baker	Back-pains	Catholic	D66	Shop-lifting

Definitions

Processing

- **ANY operation** or set of operations that is performed on personal data

Controller

- The natural/legal person or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data

Processor

- The natural/legal person or other body which processes personal data on behalf of the controller

Controller

Who is the controller of the personal data you work with ?

Answer in the Teams chat

Territorial Scope

The GDPR applies to

"…the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not."

This means…

As long as you are an EU institution, the GDPR will apply to any personal data you collect regardless of the nationality of the data subjects or the country where it is collected

Menti 6

Legal Basis

Why are you legally allowed to collect personal data?

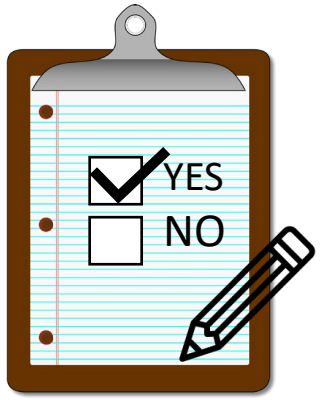
Lawfulness of Processing (Art. 6)

Personal data may only be processed if at least one of the following applies:

- Informed Consent
- Legitimate interest of the controller (DPIA)
- Public Interest
- Legal Obligation
- Contractual
- Vital interest of the data subject

Legitimate Interests of the controller

Informed consent



Informed Consent

✓Freely given

Must be a real choice and not influenced by external factors

✓Specific

Bound to several specified purposes which are sufficiently explained

✓Informed

What kind of data; How it will be used; With what purpose; Right to withdraw

✓Unambiguous

A clear affirmative statement

Informed Consent

Used as a means to meet the **Legal** and **Ethical** obligations a researcher holds towards their participants

Informed Consent: Bad examples

- What's wrong with the following statement?

	YES	NO
I do not give consent to share my data	☐	☐

Answer in the Teams chat

Informed Consent: Bad examples

- What's wrong with the following statement?

I acknowledge that the personal data collected by the researcher belongs to the university and that I have no rights to the research performed on it.

YES ☐ NO ☐

Answer in the Teams chat

Informed Consent content

Data subjects must be (at the very least) informed about:

- The controller's identity and contact details (primary contact)
- DPO's contact details (privacy@uu.nl) (secondary contact)
- Purpose of research
- Legal basis for collecting their personal data
- Personal data being collected
- Right to withdraw consent and how to do so.

Other requirements may be in place for

- Third country transfers
- Multiple controllers
- Automated Decision-making processes

Purpose Limitation

Menti 10

Purpose limitation

Why are you collecting personal data and what do you intend to do with it?

- All the data you collect has to be for a **specific purpose**
- A controllers' research objectives must be clear to the data subject

Purpose limitation and Data Reuse

The GDPR distinguishes between two types of data use:

1. Research on personal (health) data which consists in the use of data directly collected for the purpose of scientific studies ("primary use") **Initial data collection**

2. Research on personal (health) data which consists of the further processing of data initially collected for another purpose ("secondary use") **Reusing Data**

Data Reuse and GDPR

The GDPR allows for the **secondary use** of data (*further processing*) if it is for “research purposes” only if:

Appropriate technical and organizational measures are in place to ensure the privacy of the data subjects is been adequately and protected

Technical and organizational measures may include:

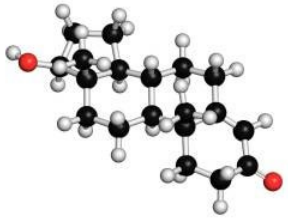
Minimization, encryption, pseudonymization, abstraction, access control …

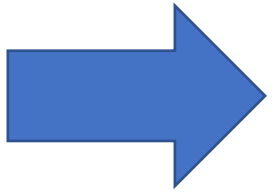
Recital 50 and Article (89)

Purpose limitation (Art. 6)

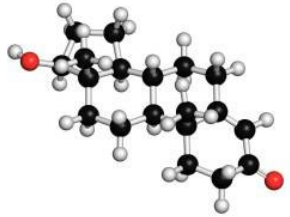
Further processing for research purposes is considered to be a **compatible purpose** (Recital 50 GDPR)

Personal Data collected for Hormone Research





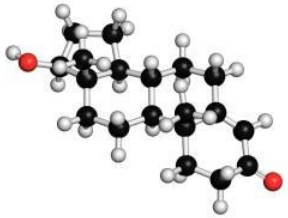
Reused for Hormone Research



Purpose limitation (Art. 6)

Further processing for research purposes is considered to be a **compatible purpose** (Recital 50 GDPR)

Personal Data collected for Hormone Research



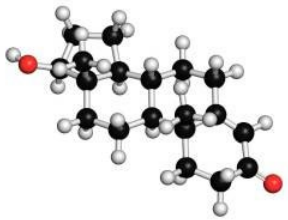
Reused for Cancer Research

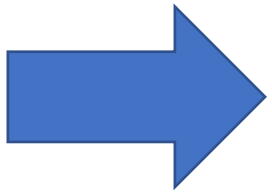
GDPR

Purpose limitation (Art. 6)

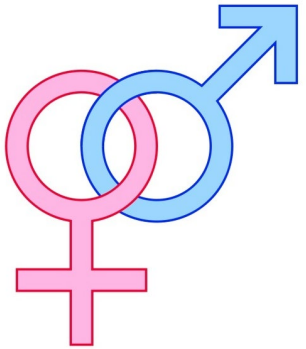
Further processing for research purposes is considered to be a **compatible purpose** (Recital 50 GDPR)

Personal Data collected for Hormone Research





Reused for Gender Studies



Ethical vs Legal

Just because it is **Legal** does not mean it is **Ethical**

Storage Limitation

Personal data should in principle not be stored for longer than is necessary.

Thankfully, the GDPR has derogation for research data:

> Personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes

Data Subject Rights

Menti 13

Data Subjects Rights

- Right of transparent communication and information
- Right of access
- Right to rectification & Right to data portability
- Right to erasure & Right to object
- Obligation to notify recipients

*Not an exhaustive list

Right to information

Data subjects must be (at the very least) provided with

- The controller's identity and contact details
- DPO's contact details
- Purpose and legal basis for collecting their personal data
- Categories of personal data
- Data Subject Rights (ability to withdraw consent, right to lodge a complaint,

Other requirements may be in place for

- Third country transfers
- Multiple controllers
- Automated Decision-making processes

Right to information

- The right to information applies regardless of the legal basis (consent, legitimate interests, etc…)
- If personal data is not obtained from the data subject, he or she must be provided the information within a reasonable period of time, but at latest after a month
- In the case that the personal data is not gathered from the data subject, in exceptional cases there is no obligation to inform.
  - This applies, if providing the information is either impossible or unreasonably expensive.
    (Meeting this requirement is exceptionally difficult)

https://gdpr-info.eu/issues/right-to-be-informed/

Right of Access

- The Data subject has the right to obtain from the controller confirmation as to whether or not personal data concerning him is being processed.
  - If so, they are entitled to gain access to such data and receive information about the data.

This can be a huge burden for researchers when not set-up appropriately

Right to rectification & to data portability

Right to Rectification

- The Data subject has the right to demand the controller to rectify, erase, amend incomplete or incorrect data

Right to Data Portability

- The Data subject has the right to transmit the data to another recipient of the data subjects choice

This can be a huge burden for researchers when not set-up appropriately

Right of Erasure

The Right to be Forgotten

- Data subject may have their personal data erased upon request
- When consent is the only legal basis for processing, the Right of Erasure is in fact the same as withdrawing consent
- When legitimate interests is the legal basis for processing, Data subjects also have the **Right to lodge a complaint.**
  (Dispute whether you could indeed have used this legal basis)

Data Management and Data Subject Rights

- RDM is necessary to comply to your Data Subject's Rights
  - Can you find individuals with ease?
  - Is rectification only necessary in one file or must it be rectified on various separate datasets?
  - Are you aware of all the locations where your data is stored? (is it still in a mobile phone? Laptop used for recording?)
  - Do you have access or know how to obtain access to all datasets if needed? (not in personal folders/laptops)

Processing Data outside of the EU

- The GDPR prohibits the transfer of personal data to countries outside of the EU unless they offer an “adequate level of protection” as determined by the European Commission.

Andorra, Argentina, Canada (commercial organizations)Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay

- A controller also may transfer personal data to a third country if it has
  - Implemented Binding Corporate Rules **(If multinational institution)**
  - Implemented standard contractual clauses **(requires legal advice)**
  - If the data subject has provided explicit consent after being informed of the risks related to the transfer (Article 46(2); Article 49(1)(a))

Processing Data outside of the EU: Alternatives

Federated Analysis:

If sufficient metadata about a dataset is present and if the dataset is in an easy-to-read format. Analysis scripts can be sent to the controllers to run the script and then send back only aggregated (anonymized data).

Hiring External:

Bring the external researcher to the data instead of bringing the data to the researcher. Highly dependent on institutional employment practices and researcher flexibility.

Synthetic Data

Sensitive Personal Data

Menti 14

Sensitive Personal Data

Personal information revealing a data subjects' ethnic origin, political opinions, religious beliefs, sexual orientation and genetic data or biometric data uniquely identifying a data subject.

In other words any interesting information about your subject…

Sensitive Personal Data

Which of the following are sensitive personal data ?

- Hair Length
- Fingerprints
- Conservative or Liberal
- Name
- Sex
- Sexual Preference

Write those that apply in Teams chat

Processing of Sensitive Personal Data

Processing of Personal data is prohibited unless:

- the data subject has given explicit consent to the processing of those personal data for one or more specified purposes
- processing relates to personal data which are manifestly made public by the data subject
- processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89

DPIA or (GBEB in dutch)

Data Protection Impact Assesment:

When: This is mandatory if data processing is likely to pose a high privacy risk for the data subjects

Recommended when unsure as to the risks.

What: During a Data Protection Impact Assessment (DPIA) you fill in a form which helps you to assess privacy issues and resulting measures to fix possible privacy problems in an early stage

DPIA breakdown I

Summary: Why are we processing this data, what is the purpose?

Legal Basis: Why are we allowed to process this data ? Informed consent?

Personal Data: Granular detail of all personal data collected

DPIA breakdown II (Privacy)

Privacy Risks: What are the risks that this personal data poses to the data subjects?

Privacy measures: What measures have been taken to protect privacy?

- Minimization?
- Separation?
- Aggregation?
- Pseudonymization?
- Privacy policy enforcement?
- Enactment of Data subject's rights

DPIA breakdown (Information Security)

Security Measures: What measures do we take to protect the data

- Security Checks?
- Access control?
- Technical safeguards?
- Encryption?

DPIA breakdown IV

Summary: Given all the information provided:

- Do we have a good reason to process the personal data
- Do we expect to be able to do so with acceptable risks to the rights of data subject?

DPIA templates

- Privacy Impact Assessment by NOREA (in Dutch, very elaborate)

- Privacy Impact Assessment for Utrecht University
  (requires logging in with your Solis-ID, based on a DPIA issued by SURF)

Security levels

Security of data files

- Encryption
- Anonymization
- Restricted access
- Secure transport
- Complete deletion

Security of computer system

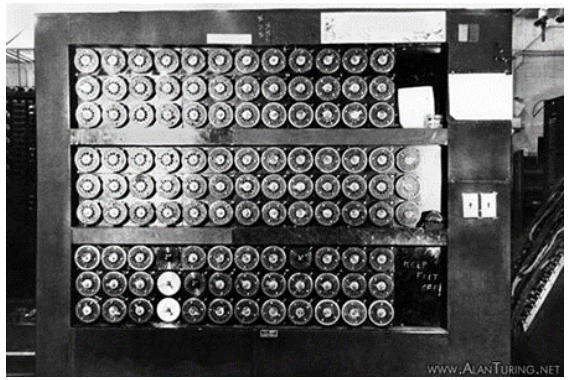
- Firewall
- Antivirus software
- Installing updates
- Using secured WIFI
- Password-protected
- Device encryption

Physical data security

- Key & Lock
- Don’t leave unattended
- Safe transport

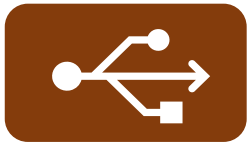
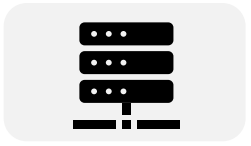
Security of data files

Encryption



Pseudonymization

Anonymization

Storage Option					YODA	DataverseNL	OneDrive	SURF DRIVE
Storage size	Varied	Varied	Varied	Varied	Varied	5GB	1TB	250GB
Price	NA	NA	Faculty	Faculty	TB €4/m	UU	UU	UU
Back-up	✗	✗	✓	✓	✓	✓	✓	✓
Controlled by UU	✗	✗	✓	✓	✓	✓	✓	✓
Internal collaboration	✗	✗	✗	✓	✓	✓	✓	✓
External Collaboration	✗	✗	✗	✗	✓	✓	✓	✓
Sensitive Information	✗	✗	✓	✓	✓	✓	✓	✓
Remote Access	✗	✗	✓	✓	✓	✓	✓	✓

Encryption

Encryption makes data unreadable/inaccessible to those without a password

Different levels of encryption

- Data encryption – the **data file** itself is encrypted
- Drive encryption – the **hard drive,** where the data file is located, is encrypted

Encryption tools

VeraCrypt

FileVault

BitLocker

Recommended

boxcryptor

https://intranet.uu.nl/en/information-security-how-do-i-handle-usb-sticks-and-portable-hard-drives

Pseudonymization

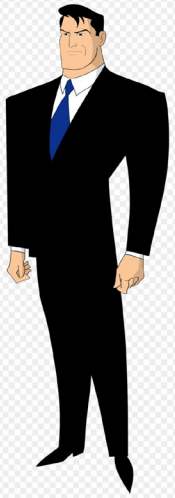
Is pseudonymized data still personal data ?

Answer in the Teams chat

Pseudonymization and Separation

Replacing personal identifiers within a database and replacing them with artificial pseudonym

The pseudonym-personal identifier relation is stored securely and separately from the pseudonymized database



Bruce Wayne → Batman

6’2 → 6’2

95kg → 95kg

1007 Mountian drive Gotham → X0978512

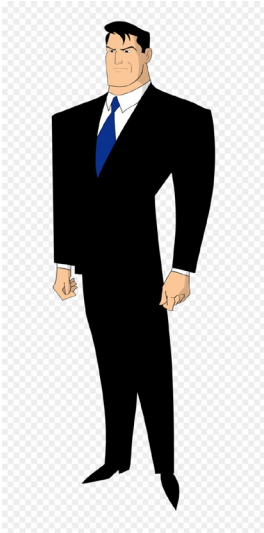
Stored Separately

Batman = Bruce Wayne

X0978512 = 1007 Mountain drive Gotham

Minimization

Removing personal identifiers within a database

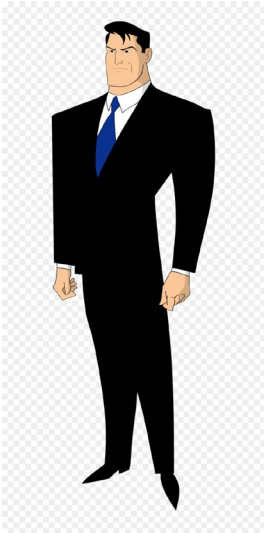


Bruce Wayne	→	Batman
6’2	→	NA
95kg	→	95kg
1007 Mountian drive Gotham	→	X0978512

Aggregation/ Abstraction

Aggregating personal data within larger bins

Reducing granularity



Bruce Wayne → Batman

6’2 → NA

95kg → Over 80Kg

1007 Mountain drive Gotham → X0978512

Anonymization

If I only published the non-identifying information of a personal data dataset.

Is the published information considered anonymous ?

Answer in the Teams chat

Anonymization

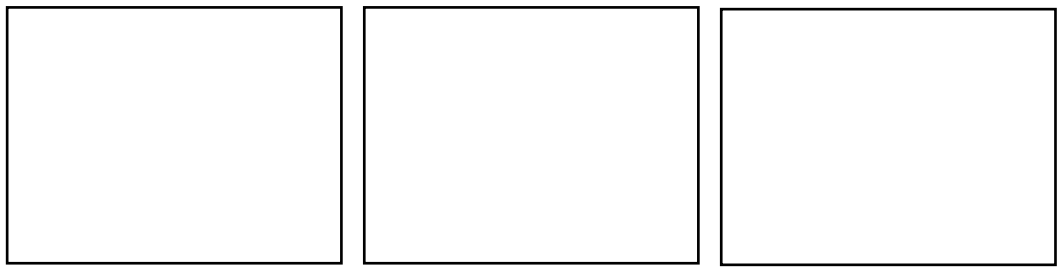
The complete and utter removal of all personal identifiers in a database.

Anonymized data can no longer be attributed to any particular individual by any means

Truly anonymized data is no longer personal data and thus no longer subject to the GDPR

True Anonymization is difficult to achieve

- True anonymization is difficult to achieve with the GDPR
- Data is anonymous only when it cannot be identified by any means “reasonably likely to be used ... either by the controller or by another person” (Recital 26).
- **Thus, even if a researcher no longer has the ability to re-identify a data set, such data set may still be regulated under the GDPR if it could be re-identified with reasonable effort by another dataset**
- Even if that dataset is unknown to the researcher



Restricted Access

- During the research
- After the research

- Who should have access to your data?
- How will you restrict access?
- How will you enable access to those authorized?
- Where will you describe who gets to access the data?

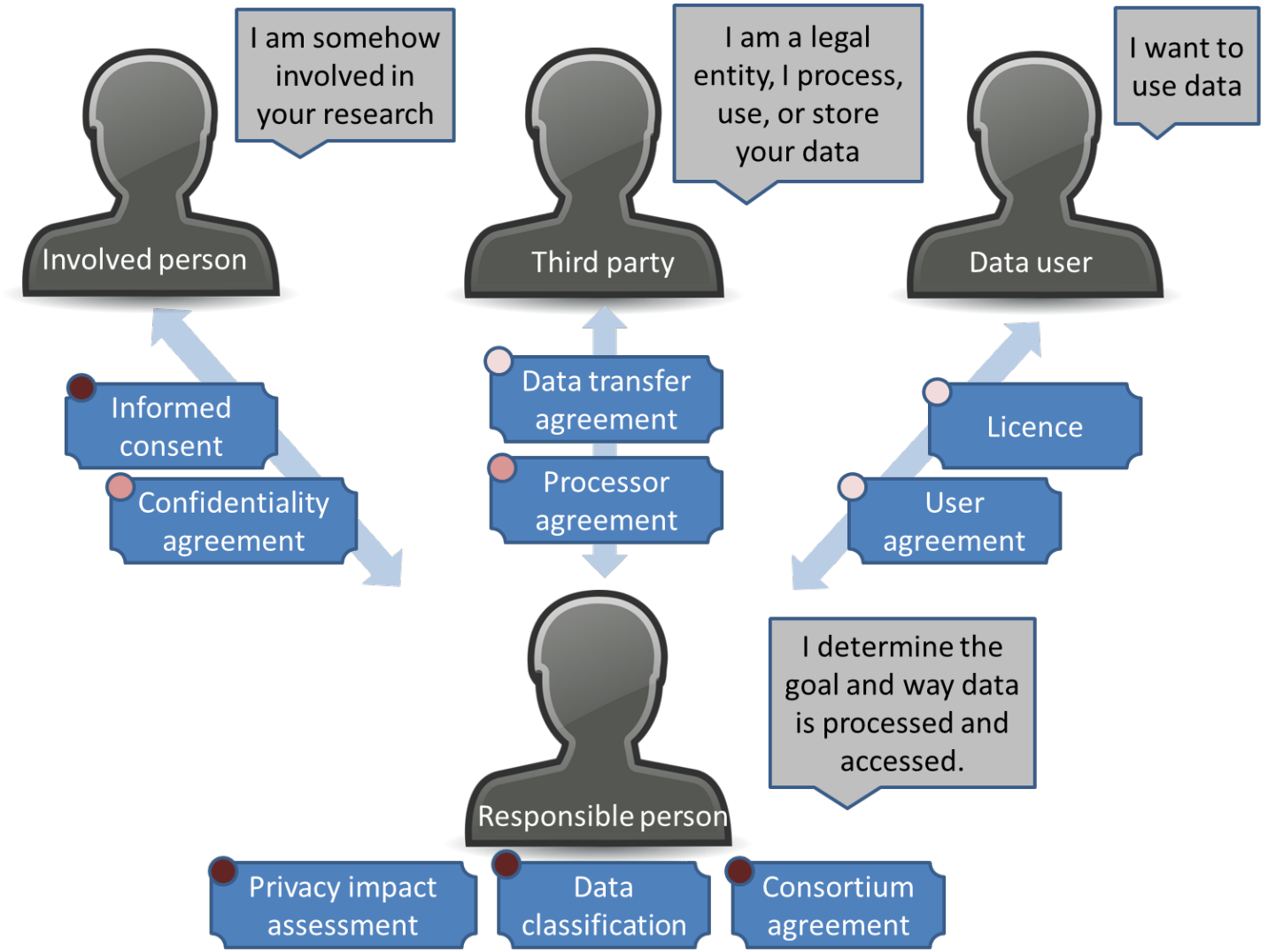
Mobile Phones and personal data

Mobile devices

- pincode -> 6 digits (non-trivial) usability features are fine, such as fingerprint
- lockout after 10 failed attempts
- android: turn on device encryption
- screen lockout after inactivity
- install only from trusted sources (Play/App store)
- do not root/jailbreak your device

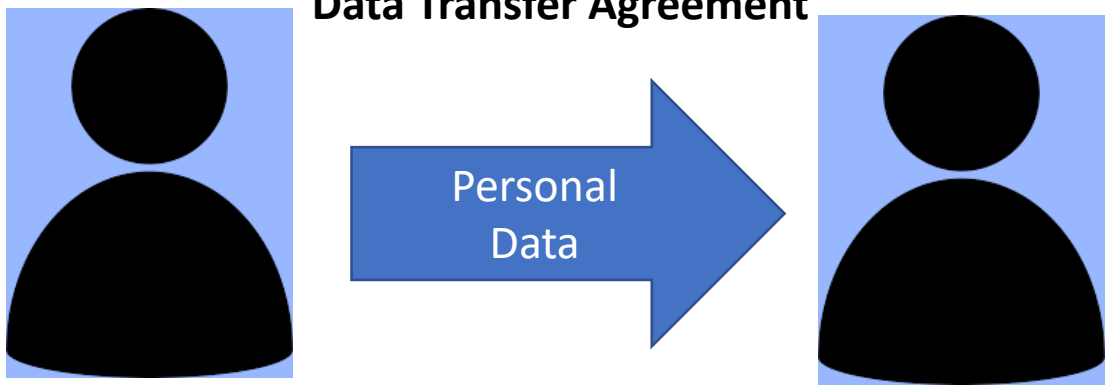
If you lose device (also personal devices with UU data!), CERT@uu.nl, they will help + you can wipe your phone remotely yourself

Legal Documents



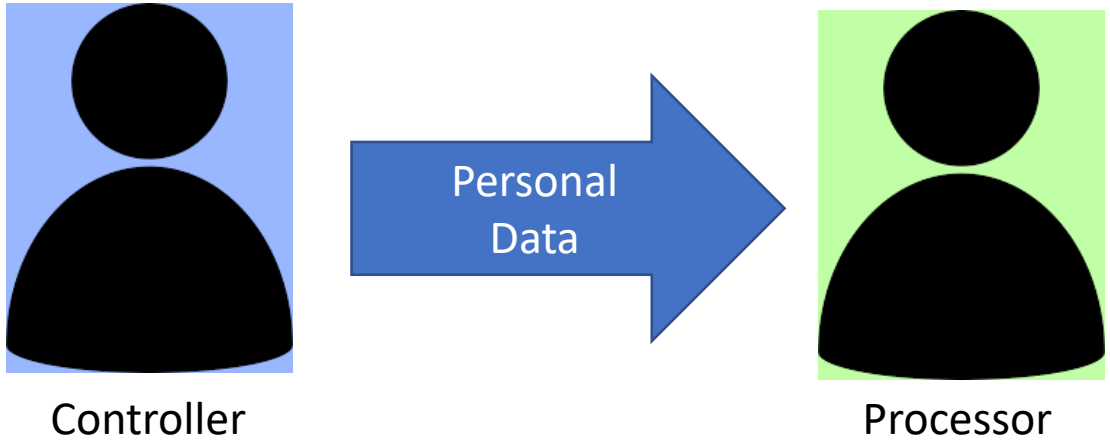
Data Agreements

Data Transfer Agreement



- Restrictions on the data processing
- Who has access
- Purposes of processing
- Responsibilities and compliance measures

Data Processing Agreement



- Exact description of how data will be processed
- Who has access
- Purposes of processing
- Responsibilities and compliance measures

Data Agreements Templates

Data Transfer Agreement

- Data Transfer Agreement as used by the YOUth Cohort study of Utrecht University.

Data Processing Agreement

- UU processor agreement
- SURF processor agreement.

Grant and Consortium agreements

- Currently lack proper safeguards and measures to protect personal data

- Should determine
    - Who are the controllers?
    - How to transfer data
    - Joint controllers
    - How data subjects can enact their rights (To object, to withdraw consent, rectification)
    - Responsibilities to the data subjects

DPO

- Work towards the compliance with all relevant data protection laws
- Monitor specific processes such as:
  - Data protection impact assessments (DPIA)
  - Employee awareness of data protection
  - Collaborating with the supervisory authorities
- Can be consulted for issues regarding the safety of personal Data
- Data subjects may contact the DPO directly for issues concerning their personal data

UU DPO: Artan Jacquet privacy@uu.nl
Edo Hoogervorst (Assistant DPO) privacy@uu.nl

UMCU DPO: Bart van Rijn privacy@umcutrecht.nl

Social Media

Sources:

Research Ethics Policy Note no. 14 RESEARCH INVOLVING SOCIAL MEDIA DATA
University of Sheffield

Social media, personal data and research guidance.
London School of Economics

Menti 17

Social Media

Is data gathered from Social Media exempt form the GDPR?

Answer in the Teams chat

Social Media

Data on social media is public but not necessarily meant for public domain

Users of social media may on one level know that they are posting to the whole world, but really only expect their audience to be friends or acquaintances they interact with on the service.

How this is perceived by the **data subject (USER)** and not the researcher -partly determines whether consent is necessary

- Users of a ‘private’ Facebook group might reasonably expect that their posts are only visible to a restricted number of people and are therefore not ‘public’ – to enter the group without the knowledge or consent of moderators and/or users would be deception

- Users of a public discussion forum on a topic with limited general interest may reasonably expect that only a small number of people are likely to view the posts – they therefore may not perceive them as public

Social Media

Consent is strongly recommended and, in many cases, necessary

- Just because personal data is in the public domain does not mean that individuals expect it to be used for another purpose like wealth screening or research
- Ethical vs Legal
- Every social media platform has different policies regarding consent procedures
  - Facebook: consent required
  - Twitter: Not required
  - Instagram: Not required
  - YouTube: consent required

Keep in mind that not requiring consent to by the social media provider does not mean you do not need consent according to the GDPR or vice-versa

Sharing Personal Data

Menti 10

“DO’S” of Sharing Data and Informed Consent

- ❑ Provide information on the intent to share the data and the conditions for sharing

Make it clear to the participant [in the information section] that one of the goals is to share the data collected with the research community.

i.e. Other researchers may request access to data in the future. Access will only be granted if they agree to preserve the confidentiality of the information as requested in this form. Their access will also require approval from the original research team.

“DO’S” of Sharing Data and Informed Consent

- ❑ Be transparent about which information you will make available

Be granular about which data will be deposited

I give permission to deposit my **impulsivity test scores, weight, age and gender data** in a repository ☐ ☐

“DO’S” of Sharing Data and Informed Consent

- ❑State the methods you will apply to reduce the risks of identification

Be specific about the methods employed to improve security and privacy

i.e. I give permission to deposit my **pseudonymized** impulsivity test scores, weight, age and gender data in a… ❑ ❑

i.e. The principal investigator will keep a link that identifies you to your coded information, but this link will be kept secure and available only to the principal investigator or selected members of the research team. Any information that can directly identify you will remain confidential. Your age and weight will be grouped into ranges (i.e. 20-30yo, 60-70kg) to reduce the risk of re-identification.

Informed Consent: Sharing Data

DON'TS

“DON’TS” of Sharing Data and Informed Consent

- Avoid terms such as fully anonymous

  Very difficult to achieve

To be truly anonymous, it should not be possible to re-identify an individual by any means. Including using external databases, even if such databases are unknown to the researcher.

“DON’TS” of Sharing Data and Informed Consent

- ❑ Avoid promises to destroy all the data

Unless absolutely certain it will be done

Have good reasons for destroying data such as

- The information has been transcribed (audio files)
- No longer needed for verification and re-use no longer expected

Be specific about which data you plan to destroy

“DON’TS” of Sharing Data and Informed Consent

- ❑ Avoid promises that all the data will only be accessed by the research team

  Instead describe explicitly which parts of the data will indeed only be accessed by the research teams and which will be available to others (after proper measures are taken to increase privacy).

How to Share personal data

In many cases the personal data you are working with is simply not fit for sharing

Other Options:

- Share the metadata and place the data under restricted access
  - When requested for the data only share it if requesters fill out a Data transfer agreement and meet the legal requirements
  - Only allow requesters to work with the data within your own facilities

Key points

- The GDPR asks researchers to be transparent towards their participants as to how their data will be handled and for what purpose.
- Personal data collected for research purposes holds a privileged spot within the legislation which softens restrictions so long as proper safeguard and measures are adopted.
- GDPR forces researchers to think about the future of the data prior to collection. This should be formulated in Grant agreements, consortium agreements, informed consent forms, DMPs and DPIAs.

Services and solutions to make research data management work

Research Data Management Support

CONTACT US >

One point of contact: info.rdm@uu.nl www.uu.nl/rdm

A network of expertise on general RDM issues, IT solutions, privacy and information security as well as legal and ethical issues