



Data Protection for Research

Diana-Leigh Watt
Information Governance Manager
D.Watt@napier.ac.uk

Everyone is responsible for data protection (in a similar way to H&S)

The University has a [Data Protection Policy Statement](#) that applies to all employees and PG Research Students. PG Research Students must sign an 'Oath of Confidentiality' to be kept by their research supervisor/supervisory team.

Data Protection is everyone's responsibility – Health and Safety style.

More information is available at: [Research, Processing Data for \(napier.ac.uk\)](#)

The DMP and the pre-DPIA

Why do we need to consider privacy and the protection of personal data for research?

What do the ICO ask for when reporting a breach?

Definition of personal data

- “Any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”
- Includes opinions

Categories of Personal Data (Standard)

The following is a list of standard descriptions of categories of personal data examples:

- Personal details, including any information that identifies the data subject and their personal characteristics, including: name, address, contact details, age, date of birth, sex, and physical description.
 - Family, lifestyle and social circumstances, including any information relating to the family of the data subject and the data subject's lifestyle and social circumstances, including current marriage and partnerships, marital history, details of family and other household members, habits, housing, travel details, leisure activities, and membership of charitable or voluntary organisations.
 - Education and training details, including information which relates to the education and any professional training of the data subject, including academic records, qualifications, skills, training records, professional expertise, student and pupil records.
 - Employment details, including information relating to the employment of the data subject, including employment and career history, recruitment and termination details, attendance records, health and safety records, performance appraisals, training records, and security records.
 - Financial details, including information relating to the financial affairs of the data subject, including income, salary, assets and investments, payments, creditworthiness, loans, benefits, grants, insurance details, and pension information.
- etc.

Special Category (Sensitive) Personal Data

Special category (sensitive) personal data concerns, reveals or is about:

- racial or ethnic origin
- political opinions
- religious or philosophical beliefs
- trade union membership
- genetic data
- biometric data (if used to identify a natural person)
- health
- sex life or sexual orientation
- criminal convictions and offences

There are also other types of data considered 'sensitive' or confidential in a research context e.g. information on rare or endangered species, commercially sensitive data, data which poses a threat to national security or would have a negative public impact – these should not be confused with special category personal data, but applying the same safeguards as you would to personal data is advisable to mitigate any risks.

Personal Data v Research Data

- Instrument measurements.
- Experimental observations.
- Still images, video and audio.
- Text documents, spreadsheets, databases.
- Quantitative data (e.g. household survey data).
- Survey results and interview transcripts.
- Simulation data, models & software.
- Slides, artefacts, specimens, samples.
- Sketches, diaries, lab notebooks

What constitutes research data?

'Research data' refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion or calculation.

*In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. **The focus is on research data that is available in digital form.***

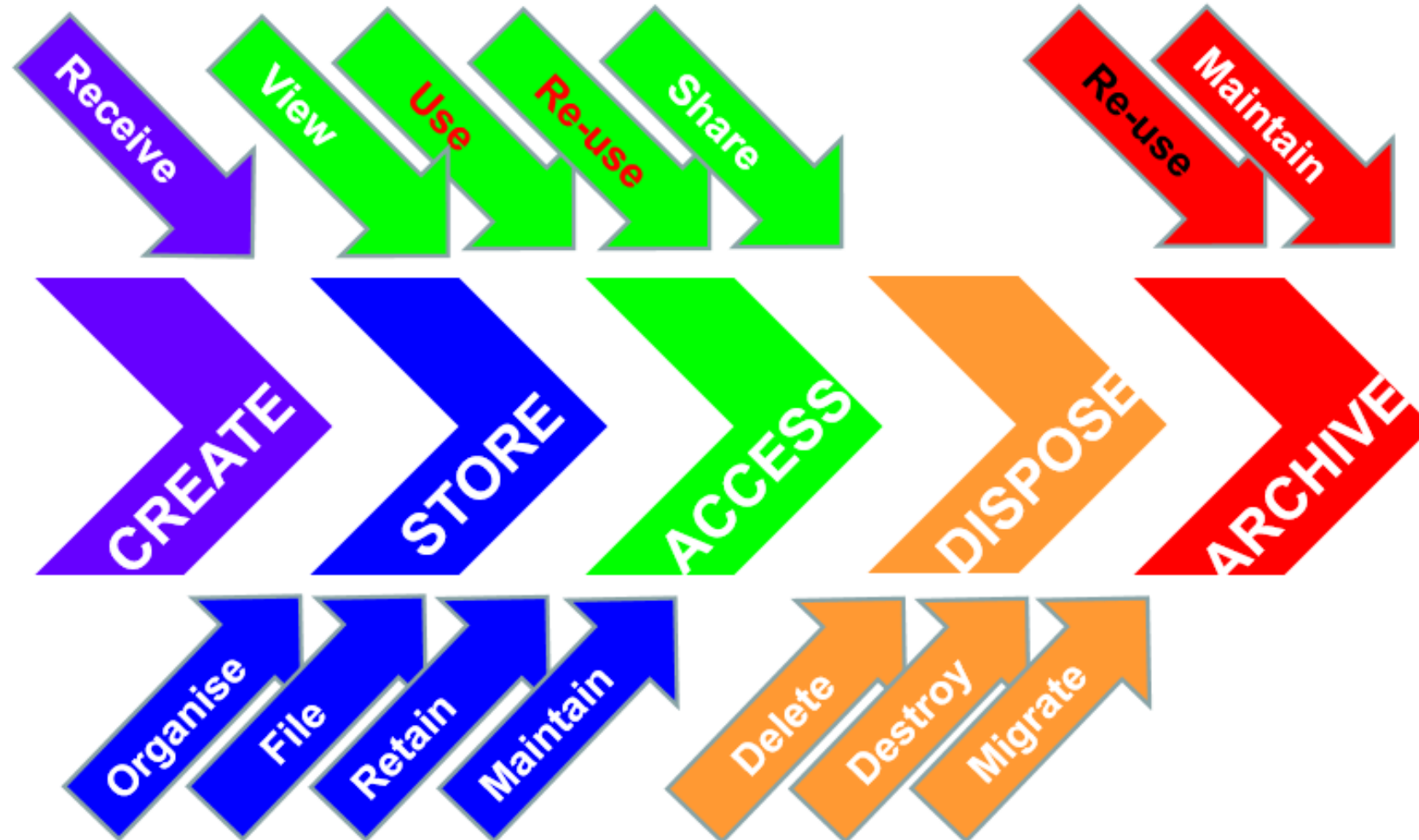
(Source: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)



What is “processing”?



Processing includes:



The Principles

- 1) Processing must be fair, lawful and transparent
- 2) Personal data collected for specified, explicit and legitimate purposes and not further processed for an incompatible purpose
- 3) Personal data must be adequate, relevant and limited
- 4) Data must be accurate
- 5) Personal data must be kept no longer than necessary for the purposes collected
- 6) Appropriate technical and organisational measures must be used to protect personal data (security)
- 7) Accountability – must be able to *demonstrate compliance*, e.g. provide records documenting processing



The Research Process

- Recruitment*
- Collecting research consent and initial personal data
- Pseudonymisation*
- Collecting personal data via the research instruments
- Security
- Retention & disposal

Participant Recruitment



- Self selecting (✓)
- v
- Targeted selection

- Targeted selection carries a higher risk of complaint or breach

Pseudonymisation vs Anonymisation

Anonymisation is the complete removal of any identifiers, which means irreversibly preventing the identification of the individual to whom the data relate. True anonymization can be complex – see the ICO’s guidance: <https://ico.org.uk/media/1061/anonymisation-code.pdf>

The individual will not even be identifiable anymore when linked with other information which is available or likely to be available.

Pseudonymisation is replacing any identifying characteristics of data with a pseudonym, such as replacing a name with something else – such as another name of the same culture for qualitative data, or a random case ID for other research.

The difference to complete anonymisation is that there will be a key to re-identify the individuals, which will be kept separately.



<https://understandingpatientdata.org.uk/what-does-anonymised-mean>

'Identifiability spectrum' by Understanding Patient Data (CC BY license)

Pseudonymisation (basic)

Password protected spreadsheet (master key) held separately to research data

The image shows a Microsoft Excel spreadsheet with a table of participant data. The spreadsheet is titled 'Repstor affinity' and has a formula bar showing '=H2'. The table has columns for Participant Number, Name, Contact Email, Age, Occupation, Gender, Ethnicity, Refugee Status, Home Address, Income, Police/Armed Forces, Benefits Recipient, and Criminal Convictions. The data rows are numbered 1 to 10. A Windows File Explorer window is overlaid on the bottom right of the spreadsheet, showing a folder named 'Project 1234 - Interviews'. The folder contains seven text files named 'Participant 001.txt' through 'Participant 007.txt', all with a date modified of 15/06/2022 12:23 and a size of 0 KB.

Participant Number	Name	Contact Email	Age	Occupation	Gender	Ethnicity	Refugee Status	Home Address	Income	Police/Armed Forces	Benefits Recipient	Criminal Convictions
001	John Doe	jdoo@gmail.com	24	Journalist	M	White-other						
002	Pamela Smith	pamela@smit.com	45	CEO	F	White-British						
003	Richard Birch	rb@outlook.com	27	Mechanic	M	Not given						
004	Sue Jones	susanjones@t.com	30	Administrator	F	BAME						
005	Isabelle Malone	dinky@gmail.com	25	Nurse	F	Asian-other						
006	Fiona Scott	f.scott69@go.com	52	Teacher	F							
007	J Bond	shaken@stirre.com	55	Not given	M							
008	Bill Johnstone	johnstone.b@com	60	Engineer	M							
009												
010												

Name	Date modified	Type	Size
Participant 001.txt	15/06/2022 12:23	Text Document	0 KB
Participant 002.txt	15/06/2022 12:23	Text Document	0 KB
Participant 003.txt	15/06/2022 12:23	Text Document	0 KB
Participant 004.txt	15/06/2022 12:23	Text Document	0 KB
Participant 005.txt	15/06/2022 12:23	Text Document	0 KB
Participant 006.txt	15/06/2022 12:23	Text Document	0 KB
Participant 007.txt	15/06/2022 12:23	Text Document	0 KB

Security

- You must have appropriate organisational and technical measures (security) in place, including assessing the security of your work environment and the electronic and manual (e.g. paper, samples, etc.) systems you use, to protect and secure personal data during processing.
- The legislation requires all processing, in transit (when being sent between entities or systems) and at rest (when in storage in a data repository or system) to be to be secure e.g. **electronic data must be encrypted at all times**. Guidance for both electronic and physical data security can be found here: [Security of Personal Data \(napier.ac.uk\)](http://napier.ac.uk). Email is NOT a secure method of transferring /sending personal data, unless encrypted, see: [Email Encryption \(napier.ac.uk\)](http://napier.ac.uk).
- Physical documents (paper) or assets (encrypted audio recorder, etc.) security in transit and storage.
- Use University approved and provided electronic systems /Apps.
- Audiovisual recorded interviews should be stored as MP3 files unless the 'visual' aspect of the recording is necessary for the project e.g. gesture/expression analysis. If using an external transcription service only MP3 files should be provided

Retention and Deletion

In relation to the 'raw' personal data (personally identifying), typically the pseudonym database/spreadsheet, voice/image recordings and non-anonymised transcripts would have a shorter retention period – probably until the findings of the study have been audited/checked/reviewed for verification purposes, but the legislation allows personal data for research to be kept for longer periods than it would for other types of processing.

Consent records would typically be kept for 6 years post-study, or longer if it is necessary to keep the personal data collected for the research for a longer period (with appropriate rationale). Please think what personal data **actually needs** to be kept in relation to the project and the practicalities of destroying it when it reaches the end of its retention period e.g. will the data be needed for subsequent research?

Externally funded projects may have retention periods dictated by funders. If you leave the University you will need to ensure that a colleague/supervisor can destroy the personal data on the date the destruction is due.

RESPECT PERSONAL INFORMATION

Ask yourself:

- Why is this information being collected?
- What will it be used for?
- Who else will be able to access the information?

Photo

Email

Date of birth

Driver licence

Bank account



Thank you!

Edinburgh Napier 
UNIVERSITY | Governance
& Compliance

