

Research Data Discovery Workshop

INTRODUCTION

The **Research Data Discovery Workshop** is for research teams.

It's designed to help us understand what data we produce in our work, and how we manage and take care of it.

INTRODUCTION

It's important that we take time to understand this.

As Researchers, we have a duty of care to accurately let people know why we need their information and what we need it for.

INTRODUCTION

It also helps us understand and prevent any risk of information getting into the wrong hands and our participants being identified.

Doing this helps us to be compliant and respect people's right to be forgotten under GDPR.

OUTCOMES

With the information from today, we'll be able to:

- Understand any risks in our process
- Think about how we might improve them
- Write better informed consent forms

How it works

HOW IT WORKS

We are going to use the **Data Discovery Cards** to help us map out the flow of data through our research process.

The workshop will take around 90 mins.

You will need a pen and the printed cards to write on.

It's not about if we are doing it right or wrong. We are only trying to understand how we currently manage data as a community.

How do we capture and manage research data?

1. Get into groups of no more than **3 people**.
2. In each group, decide on a single research method to study.
(eg. interview or usability test).
3. Write your method on a post it note as a title.
4. Think about how you do this method personally and the steps you take, starting with how information is captured in the first place.



5. Create a **Device card** to describe each device that captures the users information. (eg: A mobile phone, laptop, notebook etc).
6. Next, think about where you ultimately store your research data and fill out a **Storage card**. (eg: Your research repository, Google Drive, Dropbox etc).
7. Somehow, the data has to get from the Device card to the Storage card. Fill out a **Transfer card** to describe how this happens, then put the cards in order: Capture > Transfer > Storage to start your map.



8. Keep working through your process from left to right. Use **Device cards** if you download data to any other devices. Don't forget the Transfer card each time data moves from one place to another!

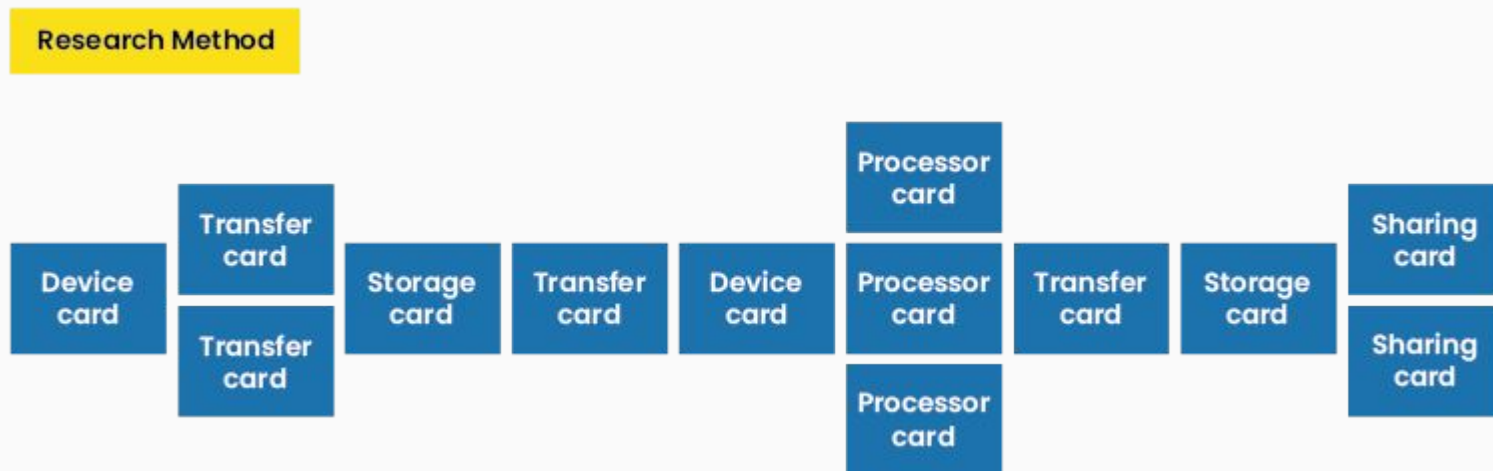
9. On your device, you might use software to edit the data. Create a **Processor card** for each tool you use if you do this.

10. If you've edited data in your process, you might want to present it to a client or the business. Fill out a **Sharing card** to describe this.

Quick check in...

CHECK IN

By now, each group should have a map that looks something like this:



CHECK IN

Next up, we're going to talk about the actual data.

First, we need to define a few terms that are important when talking about research data:

- **Personally Identifiable Information**
- **De-identification**
- **Anonymisation**

PERSONALLY IDENTIFIABLE INFORMATION

Data which relate to a **living individual** who **can be identified**, either:

a) from those data, or

b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

PERSONALLY IDENTIFIABLE INFORMATION

Some examples of **Direct Identifiers**:

- First and Last Names
- Personal identification numbers: (SSN), passport number, driver's license, etc
- Personal addresses: street address, or email
- Personal telephone number
- Photographic images (particularly of face or other identifying characteristics), fingerprints, or handwriting
- Biometric data: retina scans, voice signatures, or facial geometry
- Information identifying personally owned property: VIN number or title number
- Asset information: Internet Protocol (IP) or Media Access Control (MAC) addresses that consistently link to a particular person

Some examples of **Indirect Identifiers**:

- Date of birth
- Place of birth
- Business telephone number
- Business mailing or email address
- Race
- Religion
- Geographical indicators
- Employment information
- Medical information
- Education information
- Financial information

DE-IDENTIFICATION

A process of removing or masking **direct identifiers** in personal data such as a person's name, face or voice, address, Driving Licence or other unique number associated with them.

De-identification is sometimes called pseudonymisation.

<https://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf> (Page 15)

ANONYMISATION

A process of ensuring that the **risk of somebody being identified** in the data is negligible.

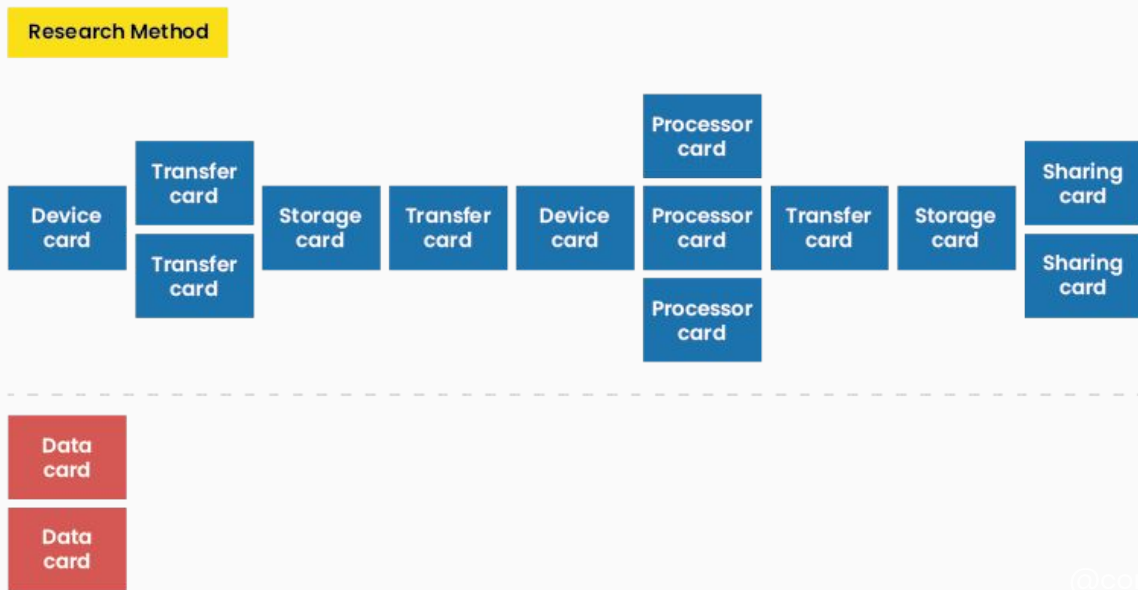
This involves doing more than simply de-identifying the data, and often requires that data be further altered or masked in some way in order to prevent statistical linkage.

<https://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf> (Page 16)

Let's get back to it

CAPTURING RESEARCH DATA

11. Next, we need to represent the data itself as it moves through our process. Under your capture Device card(s), start a new row for data.





12. Create a **Data card** for each type of data that is created at the point of capture. (eg: photos, videos, audio, notes etc)

13. Work through your process. Think about each time that data is copied, changed or added to another file. Create new **Data cards** to represent these changes.

14. At what point does the data get deleted? Is there anything left on devices or storage for longer than you actually need it?

Reflections



Staying in your groups, discuss the following:

- How did you find the exercise?
- Will you do anything to change your ways of working?
- Does your consent form reflect how you handle research data?
- Was data deleted from all devices?
- How long do you actually need to keep data for?
- Were you ever transferring data over unsecured networks?



Each group take turns to share your findings with the workshop.

Walk them through the process you have mapped out and talk about your reflections.

In particular, call out anything you felt was difficult or challenging.

Thank you!