UKAN

UK ANONYMISATION NETWORK

Table of Contents

0.1 The Anonymisation Decision-Making Framework

Anonymisation is a process that balances producing safe data with maintaining useful data.

Anonymisation is about managing risk. Anonymising data is *not* about removing all risk, but about taking precautions. When anonymisation is done well the risk of disclosing information referring to individuals should be *negligible*.

The course will take you around one hour to complete. Currently we do not offer a feature to save your progress, but we are storing the session in a browser cookie. This means your browser should remember your progress.

About

This course has been developed by the UK Anonymisation Network (UKAN) based on the Anonymisation Decision-making Framework.

Important: This course gives a taster of the things you need to consider if you want to anonymise data effectively. We recommend that you seek expert advice before releasing/sharing data derived from personal data.

UK Anonymisation Network

0.2 How to use the Anonymisation Decision-making Framework (ADF)

The ADF can be used at three levels, as:

1. A tool to boost your data knowledge
2. A guide to help you understand your data situation
3. A practical guide to help you decide how to anonymise your data

You can use the ADF in different ways depending on your knowledge and skills.

At level 3 you will need to make complex judgement calls about when data is sufficiently anonymised given your data situation. The ADF will help you to make sound decisions based on best practice. However, it can only do this within the context of the knowledge and skills you bring. Where necessary, you should seek advice from appropriate experts.

The relationship between data and its environment is referred to as a ‘data situation’. In a *dynamic* data situation, data are moved from at least one environment to another. How a data custodian has come by the data, the data itself, what role and responsibilities the data custodian(s) has and to whom, whether it shares and/or releases products of that data and how it does this, all come together to create particular data situations.

1.1 Describing the data situation

To anonymise data it's important to take into account its 'environment'.

What do we mean by an environment?

It is, in a broad sense, the context for data. An environment includes:

» other data.
» people who interact with the data
» their responsibilities and governance, and
» the infrastructure for processing and protecting it.

What do we mean by a data situation?

The relationship between data and its environment is referred to as a ‘data situation’. How a data custodian has come by the data, the data itself, what role and responsibilities the data custodian(s) has and to whom, whether it shares and/or releases products of that data and how it does this, all come together to create particular data situations.

1.2 The Anonymisation Decision-making Framework

This course is structured around the Anonymisation Decision-making Framework (ADF). You can approach the components in a flexible order because they are not a checklist. The 10 components are:

1. Describe your data situation
2. Know your data
3. Understand the use case
4. Understand the legal issues
5. Understand the issue of consent and your ethical obligations
6. Identify the processes you will need to assess disclosure risk
7. Identify the disclosure control processes that are relevant to your situation
8. Identify who your stakeholders are and plan how you will communicate

9. Plan what happens next after you have shared of released data

0. Plan what you will do if things go wrong

1.3 Describing the data situation

Below is an example of a data situation

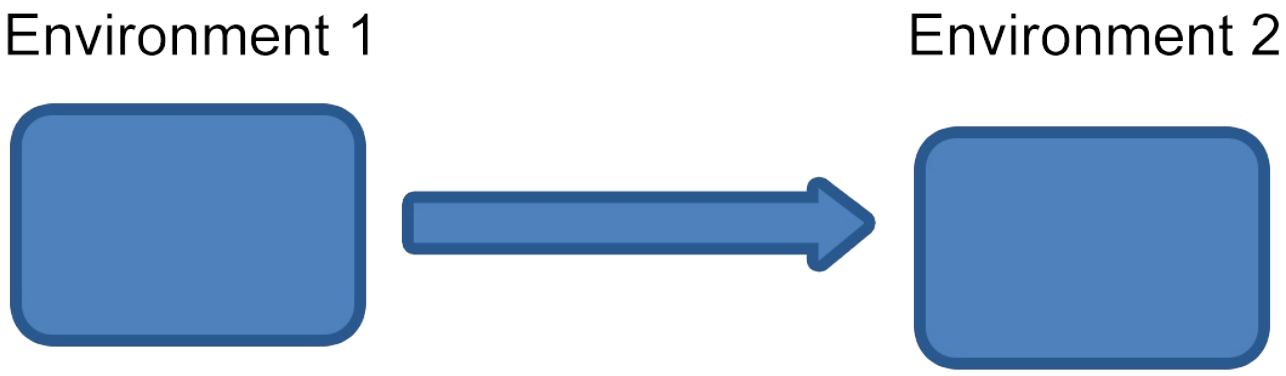
Busmapper Ltd. collects data from its customers on public transport use. It wishes to share an anonymised version of the data with the Local Authority (LA) to support the LA provision of public transport. This is environment 1.

The LA has signed a data sharing agreement with Busmapper Ltd. to analyse the data. This is environment 2.

Critical issues when describing your data situation include:

» your data
» legal and governance issues and
» the issues around consent and ethics

We will cover them in component 2, 3, 4 and 5 of the ADF.



What happens within and across the environments is the data situation.

Topic 1 Summary

1. The Anonymisation Decision-making Framework has 10 components and can be used at three levels, as:
   » A tool to boost your data knowledge
   » A guide to help you understand your data situation

» A practical guide to help you decide how to anonymise your data

2. Anonymisation is a process that balances producing safe data while maintaining useful data.

3. When anonymisation is done well, the risk is negligible, not zero.

A more detailed description of the Anonymisation Decision-making Framework can be found here: www.ukanon.net

2.1 Know your data

If you want to anonymise a dataset, it is crucial that you understand your data. For example, you have to know

» where there are special or unique cases,
» which combination of variables are risky, and
» what is sensitive information.

Let's look at an example dataset to illustrate this: the Titanic passenger data. The data describes the survival status of individual passengers on the Titanic. It has 1309 passengers (rows) and 14 variables (columns). The variables are characteristics such as name, age, travel companion or how much they paid for the ticket. An extract of the data is below.

You can find the full dataset here or here .

You can find out more details about the data if you follow this link .

Passenger class	Survived	Passenger name	Sex	Age	Fare
3	1	de Messemaeker, Mrs. Guillaume	female	36	17.400
2	0	Levy, Mr. Rene Jacques	male	36	12.875
1	1	Barkworth, Mr. Algernon	male	80	30.000
3	0	Goodwin, Miss. Lillian Amy	female	16	46.900

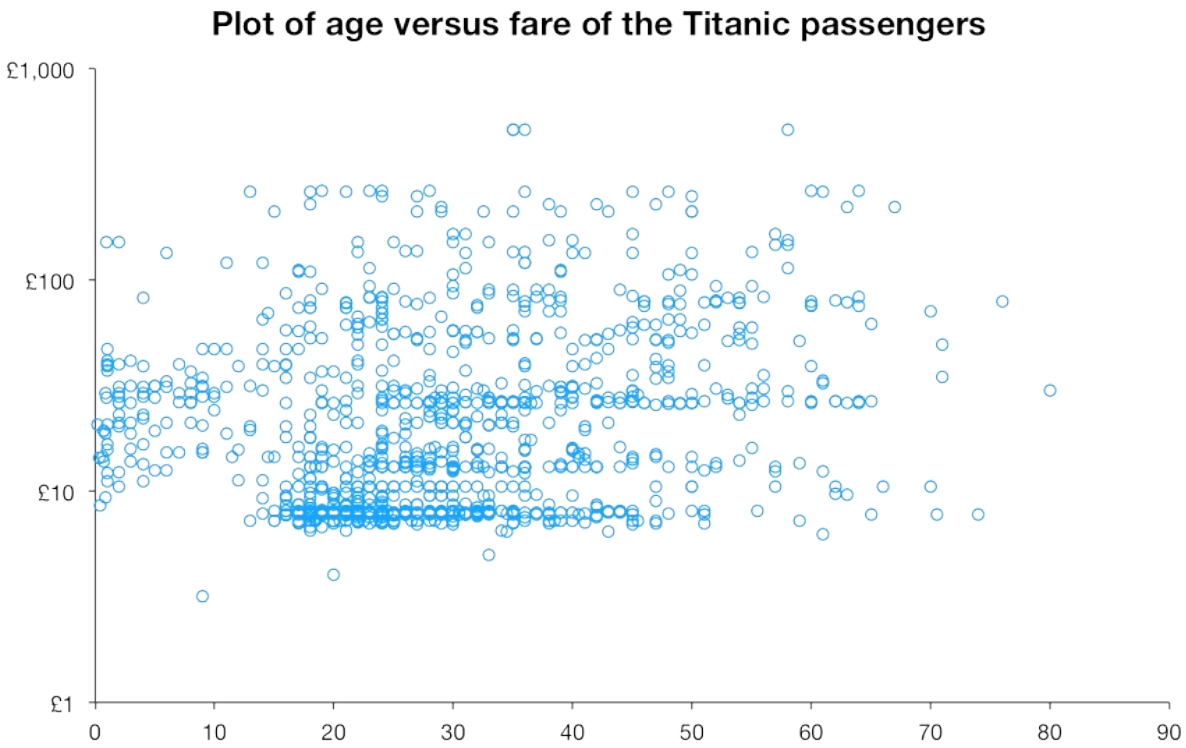
en.wikipedia.org/wiki/File:Stöwer_Titanic.jpg

2.2 Know your data

Here are some suggestions for getting to know **quantitative data** .

1. ‘Eyeball’ your data if it is small enough

2. Create descriptive statistics, for example a frequency table of passenger class



3. Create a chart, for example, a scatterplot of age versus fare

4. Create other statistics, such as a count of the missing cell values in the age column

You should do this for all variables in the data.

2.3 Know your data

You may also have **qualitative data** , for example, in the form of free text fields in a survey. They can be problematic when releasing them in its raw form because you have to be sure the data does not contain a name or other information relating to individuals.

2.4 Know your data

Below are a few selected rows and columns of the Titanic passenger data. (The full data is here: here or here).

This data needs to go through the whole anonymisation process, but there are a few basic options to change the data. **Consider which of the variables below you would expect to be relatively risky, and why.**

Passenger Class Choose Option	Survived Choose Option	Passenger Name	Sex Choose Option	Age Choose Option	Fare Choose Option
3	1	de Messemaeker, Mrs. G	female	36	17.400
2	0	Levy, Mr. Rene Jacques	male	36	12.875
1	1	Barkworth, Mr. Algernon	male	80	30.000
3	0	Goodwin, Miss. Lillian An	female	16	46.900

Just applying some anonymisation method on its own is not enough. The beginning of the process of anonymisation is to understand your data.

Passenger class

This variable is most likely less risky than others. However, we should not ignore it completely, for example, in combinations with other variables.

Survived

This variable is most likely less risky than others. However, we should not ignore it completely, for example, in combinations with other variables.

Passenger name

If a name is unique, we would call it a direct identifier . This means an unusual name is enough to identify a person. A more common name such as James Smith, in combination with an address, for example, may also be a direct identifier. In general, you have to remove direct identifiers.

Ticket number (in the full dataset) may also be a direct identifier. If someone knows who the ticket number refers to, it would be possible to identify the person in the data.

What can we do? A common strategy is to mask or hash the direct identifier, that is, to replace it with another, usually random, value. Or we may choose to simply delete both columns.

Sex

This variable is most likely less risky than others. However, we should not ignore it

completely, for example, in combinations with other variables.

Age

Age is quite an important variable for many analyses or use cases. A common strategy is to create age bands such as 0-10, 11-20, 21-30 and so on. Choosing the optimal trade-off between usefulness and risk is up to the analyst and also depends on the remaining data and its context.

Fare

Fare is, similar to age, quite granular. We may want to include it in a data release because it could tell an interesting story around prices. Again, it may be prudent to reduce the amount of information disclosed. More on possible methods to do so in topic 7 of the course.

Topic 2 Summary

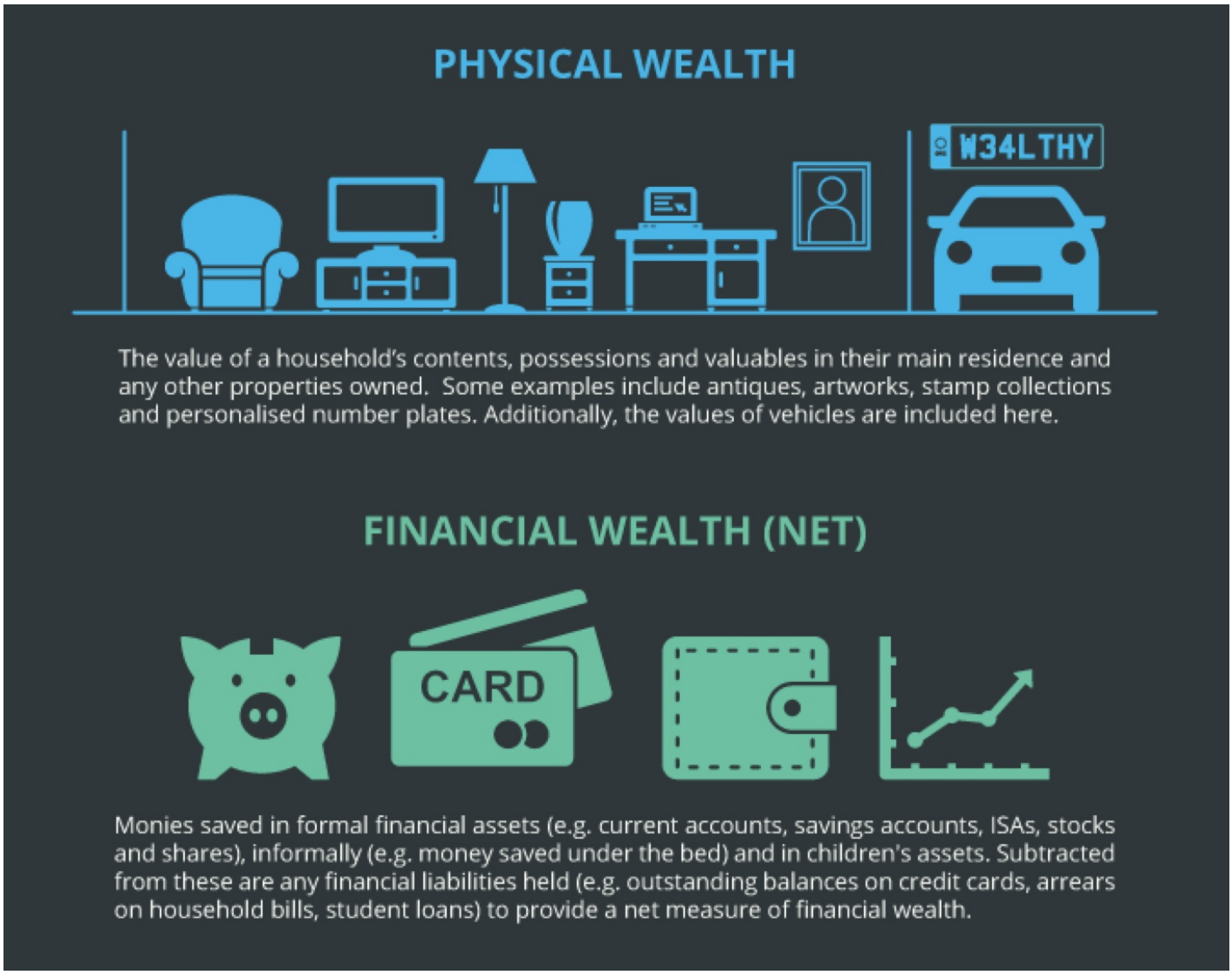
1. To anonymise your data, you have to understand it. For example, understanding what are special cases, outliers, combinations of variables, and what is sensitive information are all important steps.
2. Techniques for understanding your data range from visualisations to descriptive statistics.
3. Consider carefully which variables may pose a risk to creating safe data.

3.1 Understand the use case

In determining the use case for your data you should capture three issues, they are:

1. Why you wish to share or release your data
2. Who will want to access your data
3. How those accessing your data may want to use it

Your use case provides context on how and what data you should release. If your core users want access to a particular variable you consider risky, you may need to restrict who has access to that data, via accredited researchers, for example. To meet your obligations to produce *safe, useful data* you may also share it with a wider audience under a different licence by restricting a lot of the detail.



For example the [Wealth and Asset Survey] looks at personal **financial information**. These variables are usually anonymised or even omitted. But in this case they are the core variables,

so the anonymisation methods focus on other information like demographics.

Topic 3 Summary

1. When finding your use case, capture three issues
    - » Why you wish to share or release your data
    - » Who will want to access your data
    - » How those accessing your data may want to use it
2. Your use case provides context on how and what data you should release.
3. There are a range of methods to gain information such as workshops, consultations or user feedback.

4.1 Understand the pre-share/release legal issues

Before you anonymise your data you will be processing personal data and therefore will need to understand your responsibilities under the Data Protection Act (1998).

What is the Data Protection Act (1998)?

The Data Protection Act covers the processing of data of identifiable living individuals (personal data) and requires consideration of "likelihood of identification". The DPA does not require anonymisation to be completely risk free – the risk of identification should be negligible.

In the UK some courts use the "likely, reasonably" test. You should make all reasonable attempts to limit the likelihood of identification, where necessary seek advice from experts such as UKAN and/or seek appropriate legal advice. Ensure this process is documented and kept up-to-date. In this course we recommend UKAN's anonymisation decision-making framework.

For further information please see

» Data Protection Act (1998)
» ICO Guide to Data Protection
» UKAN Best Practice Guide: Anonymisation Decision-making Framework

Data Protection Act 1998

Data Protection Act 1998

4.2 Understand the pre-share/release legal issues

How to ensure you meet your legal responsibilities

Having clearly defined governance structures and processes in your organisation will help you follow best practice in anonymisation and meet your legal responsibilities.

Let us consider what we mean by **good governance** for your organisation:

» Ensuring someone with the suitable skills and knowledge oversees the anonymisation process
» Ensuring you have procedures for identifying cases where anonymisation may be problematic or difficult to achieve in practice
» Documenting who does what, when and how and accurately record the anonymisation process
» Conducting a privacy impact assessment
» Training all staff appropriately and update training periodically
» Knowing what to do if, in the rare event, things go wrong (discussed further on in the course)
» Ensuring you keep up-to-date with any new guidance or case law that clarifies the legal framework surrounding anonymisation

Find more information on this at the website of the Information Commissioner's Office.

Topic 4 Summary

1. Before you anonymise data, it will be personal data. This means that in the UK you have to know about the Data Protection Act (1998).
2. Good governance for your organisation means a range of steps such as documenting the anonymisation process.
3. Further information is in the Information Commissioner's Office Guide to Data Protection

5.1 Understand the issue of consent and your ethical obligations

When you collect data, it is always preferable to obtain consent from those the data relates to. Consent ought to promote both trust in your organisation and the continued support of your data subjects.

What is consent?

Consent is derived from the principle of autonomy. In the data context this translates to enabling data subjects to have some say in how data about them is used, even when it is anonymised.

Consent, as in 'informed consent', requires that it is voluntary, given in full knowledge and given by someone with capacity to consent.

Consent is sometimes confused with notification. A statement about the planned use of data may leave people no option but to agree because they cannot otherwise access the data product or service. This is not informed consent but a form of notification

When consent is not possible

Sometimes it is not possible to obtain consent for secondary use of anonymised data because for example you collected your data too long ago or the cost of doing so is prohibitive. The Data Protection Act (1998) offers provision for this situation. Section 33(2) stipulates,

For the purposes of the second data protection principle, the further processing of personal data only for research purposes in compliance with the relevant conditions is not to be regarded as incompatible with the purposes for which they were obtained.

5.2 Understand the issue of consent and your ethical obligations

The key ethical obligation is that you should do no harm.

Most professions have ethical guidelines

Professional ethics should dictate that you treat people with respect, which means doing all you can to avoid harm. Ethical guidelines and codes of conduct that apply to statisticians should also apply to data publishers. For example, the Office for National Statistics Code of Practice speaks of "serving the public good".

You should bear in mind what impact the data release might have

In practice the (anonymised) data you share or release might have an impact on the data subject or how they feel about how it is used. The ICO makes the point that even anonymised data does not mean you can always disclose it when you take into account ethical considerations. (See ICO Anonymisation: managing data protection risk code of practice).

5.3 Understand the issue of consent and your ethical obligations

Two examples that consider both legal and ethical considerations

Example 1: Releasing (personal) data about deceased persons

You can legally publish data of deceased people (unless an exemption applies). However, this should not negate ethical considerations such as any distress you may cause to the relatives of the deceased persons by releasing such data.

For further information and the list of exemptions see - Confidentiality guidance: Disclosure after a patient's death Confidentiality guidance: Disclosure after a patient's death

5.4 Understand the issue of consent and your ethical obligations

Two examples that consider both legal and ethical considerations

Example 2: Releasing anonymised data

The Data Protection Act (1998) does not prevent the disclosure of anonymised data, however there may still be reasons for not disclosing it. For example, in cases where individuals, though not identifiable, are at risk of misidentification and where this misidentification may leads to harmful consequences.

The key point is that you will need to consider both what you can do legally and what you should do ethically. Note that they may not be the same thing. Ethical considerations may override legal ones.

Topic 5 Summary

1. Consent is always preferable to obtain because it meets ethical standards and promotes trust in your organisation.
2. Your key ethical obligation is that you should do no harm.

3. You will need to consider both what you can do legally and what you should do ethically. They may not be the same.

6.1 Identify the processes you will need to assess disclosure risk

The key question is - *how might a data breach actually occur?*

We recommend that you follow step 1, assess the data in front of you, and step 2, scenario analysis.

Step 3 is both informative and good practice but takes time and resources. Therefore, we recognise that it is not an option that everyone can undertake. However, it still may be necessary – consult an expert.

The three steps are as follows:

1. Assess the data in front of you
2. Scenario analysis
3. Testing the scenarios

6.2 Identify the processes you will need to assess disclosure risk

Step 1. Assess the data in front of you – getting to know your data

This is covered in component 2 ‘Know your data’ section of the framework. To summarise you should be able to identify:

» whether it is aggregate data or data with individuals
» what the key variables are
» what is likely to affect the risk level of your data such as
  » the sensitivity of your data (health data for example)
  » the disclosiveness of your data (many variables for example)
  » who has access to the data (that is the data environment in general)

6.3 Identify the processes you will need to assess disclosure risk

Step 2. Scenario analysis – getting to know your data within the share/release environment

Apply the information from step 1 ‘Assess the data in front of you’ to evaluate if your data is at risk. There are three things to consider:

1. other relevant external data sources
2. potential intruders

3. the likelihood of an attempt of attack on dataset

This information will help you establish **relevant plausible scenarios for your data situation**. When you undertake a scenario analysis, you are essentially considering the how, who and why of a potential breach.

An example of a scenario analysis framework

» Scenario description: specify a brief summary of the event.
» Rationale: Specify what lies behind the event, i.e.: (i) malicious intent, (ii) non-compliance with data security policies and procedures, (iii) error.
» Likelihood of the scenario: specify what factor if any may influence the likelihood of the event such as easier ways of obtaining 'the data' or legal penalties or employee training and support to ensure compliance.
» Mitigation: specify what if any action you can take in order to reduce the risk of the event in terms of its likelihood or its impact.

Key question to help you develop plausible scenarios include:

1. Who might be interested in attacking your data?
2. Why are they interested?
3. Are there other, easier, ways for them to achieve their goals?
4. What other data would they need to re-identify someone in my anonymised dataset?
5. What might be the impact or consequences?
6. Into which scenarios should I invest my resources?

Finding other external data sources

Below are examples of other sources of data that you may want to consider when developing your scenarios of disclosure.

Public sources of data: This includes such as public registers, professional registers, electoral register, land registry, estate agents lists, newspaper reports, archived reports and announcements, parish records, vital statistics such as birth, death and marriage records.

Other similar data releases: This includes such as releases from your partner organisations and other organisations in your industry or sectors

Official data releases: This includes such as data releases from the Office for National Statistics, Government departments and agencies and Local Authorities.

Restricted access data sources: This includes such as social media networks and potentially any other organisation collecting data

6.4 Identify the processes you will need to assess disclosure risk

Step 3. Testing the scenarios developed in step 2 – penetration testing

The key is to document each step you take to assess and protect against risk.

Penetration testing can be useful to validate assumptions by simulating attacks using "friendly" intruders. The ICO recommend carrying out a motivated intruder test as part of a practical assessment of a dataset's risk.

» Are there other sources of data that can be linked that haven't been considered?
» How easy is it to manipulate the data and find people in practice?
» Can you gauge an approximate level of risk in the data?

Penetration testing can be useful to validate assumptions by simulating attacks using "friendly" intruders. The ICO recommend carrying out a *motivated intruder* test as part of a practical assessment of a dataset's risk. For example, the statisticians at Department of Energy & Climate Change went through several tests for the national energy efficiency data. A prize for correctly re-identifying an individual was not claimed by a group of motivated postgraduate students in computer science.

There are also statistical techniques for assessing risk although these are complex and can be difficult for a non-expert to interpret.

Topic 6 Summary

1. In practical terms, as a minimum we recommend step 1, assess the data in front of you, and step 2, scenario analysis.

2. When you undertake a scenario analysis, you are essentially considering the how, who and why of a potential breach.

3. The key is to document each step you take to assess and protect against risk.
4. You may need expert input to carry out more technical analyses.

7.1 Identify the processes you will need to control disclosure risk

Anonymising data is an iterative process. The key is to recognise that the technical work is only a part of the anonymisation process.

Every anonymisation process depends on the data situation at hand. We will show a few techniques with the help of the Titanic passenger data. Depending on the use case, you may reach a different decision, even with the same dataset.

The Titanic passenger data is the same as in component 2 ‘Know your data’.

You can find the full dataset here or here .

You can find out more details about the data if you follow this link .

en.wikipedia.org/wiki/File:Stöwer_Titanic.jpg

7.2 Identify the processes you will need to control disclosure risk

Assume our aim is to release this clearly personal dataset (contains names!) as an anonymised dataset. We’ll ignore for a moment that the passengers are long dead, that we’re unclear about the use case and any other of the 10 components.

The following is an example process. By no means it implies that this is the right, perfect or only solution.

Step 1: We have completed component 2 ‘Know your data’. If you haven’t done this yet, doing it now may help you understand the data better.

An extract of the Titanic passenger data

Passenger Class	Survived	Passenger Name	Sex	Age	Number of siblings/spouse	Fare
3	1	de Messemaeker, Mrs. Guillaume	female	36	1	17.400
2	0	Levy, Mr. Rene Jacques	male	36	0	12.875
1	1	Barkworth, Mr. Algernon	male	80	0	30.000
3	0	Goodwin, Miss. Lillian Amy	female	16	5	46.900

7.3 Identify the processes you will need to control disclosure risk

The following is an example process. By no means should it be considered to be the it implies that this is the right, perfect or only solution. What you will need to do will depend on the key properties of the data environment that the data are being moved to.

> First, we remove direct identifiers. The most obvious ones are name (column 3) and ticket number (column 8). We drop ticket number (very limited use) and replace the names with a randomly assigned number (a process called pseudonymisation). Don’t forget to keep a copy of the original data!
>
> We redact further variables that seem not useful or create too much risk. Therefore we delete cabin (column 10), lifeboat (column 12), body identification number (column 13) and home/destination (column 14).
>
> It is very common to group age, so after some experimentation we decide on the following age bands: 0-14, 15-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55 and older.
>
> A frequency table reveals that the number of siblings/spouses aboard

(column 6) has two very large families that are thus outliers. We expand the group “four” to include “four or more” siblings/spouses aboard.

A similar frequency table for number of parents/children aboard (column 7) reveals a similar risk. We group “three” to include “three or more” parents/children aboard.

The passenger fare is very granular and therefore risky. We round the variable to it’s nearest 10 and inspect it again.

We learn that home/destination is crucial for most analyses. We therefore go back to group the geographic destinations manually and add them back into the dataset.

The variables passenger class (column 1), survival status (column 2) and sex (column 4) look ok on their own. However, the may still be indirect identifiers in combination with other variables. In other words, we have to analyse the combinations of variables to understand what disclosure risk remains.

Most disclosure control techniques can be combined. We may chose to only release an aggregate table.

Many more techniques and examples are in the Information Commissioner’s Office’s document Anonymisation: managing data protection risk code of practice.

7.4 Know the processes you will need to go through to anonymise your data

BEFORE

Passenger class	Survived	Passenger name	Sex	Age	Number of siblings/spouse	Fare
3	1	de Messemaeker, Mrs. Guillaume	female	36	1	17.400
2	0	Levy, Mr. Rene Jacques	male	36	0	12.875
1	1	Barkworth, Mr. Algernon	male	80	0	30.000

3	0	Goodwin, Miss. Lillian Amy	female	16	5	46.900

AFTER

Passenger class	Survived	Passenger name	Sex	Age	Number of siblings/spouse	Fare
3	1		female	35-44	1	17
2	0		male	35-44	0	13
1	1		male	55 or older	0	30
3	0		female	15-19	4 or more	45

Topic 7 Summary

1. Anonymisation methods are only one step in the anonymisation process.
2. Every anonymisation process depends on the data situation, the use case, the environment and so forth.
3. Many more techniques and examples are in the Information Commissioner's Office's document Anonymisation: managing data protection risk code of practice.

8.1 Identify who your stakeholders are and plan how you will communicate with them

Different audiences will need different communication strategies.

» How much of your anonymisation process do you want to explain?
» Who are you informing about your data release?
» Some people may want to learn about the benefits, while others care more about your governance.

Possible stakeholders might be:

» The *data subjects*, the people linked to the data
» Funding providers
» Users and researchers, accessing the datasets
» The general public, who might be concerned about your data release
» The regulators or related organisations

8.2 Identify who your stakeholders are and plan how you will communicate with them

The communication strategy also depends on your data release, for example, who has access to the data.

Make sure you know how much you want to explain about your **anonymisation methodology**. As a guideline, the more you are comfortable to release without affecting the disclosure risk, the better. For example, you want to tell users that you have added noise to a variable and, as much as possible, how this affected the variable. Of course, you cannot disclose the exact amount of noise.

Why and how the release is useful for people should also be reflected in your communication strategy.

Topic 8 Summary

1. Different audiences will need different communication strategies.
2. Tell people as much as possible about your anonymisation process and methodology.
3. Tell people the why and how your release is useful.

9.1 Plan what happens next once you have shared or released data

Your work is not done once you have shared or released a dataset. In other words, **you should not share/release and forget.**

Any data environment is not static and over time is very likely to change. This could impact the risk level associated with your data.

A few examples of how the data environment might change include:

» you release a new, more detailed version of the data
» you release additional and related data
» someone else releases related data
» you increase access to more users
» someone provides feedback about legal or ethical aspects
» the data is used in new ways

9.2 Know what to do if things go wrong

Revisit on a regular basis data that you have shared or released. It is good practice to periodically reconsider both its risk and usefulness.

At the very least, you should:

» keep a register of all the data your share or release
» consider any new shares and releases against previous shares and releases
» scan the data environment by looking at
  » other similar data released by organisations in your sector
  » publicly available data sources
  » how and who uses your data

Topic 9 Summary

1. Your work is not done once you have shared or released a dataset.
2. The data environment is very likely to change over time.
3. This means you should revisit on a regular basis data that you have shared or released.

10.1 Plan what you will do if things go wrong

Sometimes even when you follow all the recommended advice things can go wrong. The risk of disclosure is not zero (it should be negligible).

It is important that you put in place mechanisms that can help you deal with a data breach if, in the rare event, one were to occur.

Developing a breaches policy should:

» capture what you mean by a breach
» identify critical factors that might lead to a breach such as non-compliance, errors or malicious intrusion
» outline the penalties associated with a breach
» identify who is responsible for doing what, when and how

10.2 Plan what you will do if things go wrong

Anticipating a data breach you will give you some control over its initial and long-term impact on your stakeholders and yourself.

If a breach were to occur you will need to:

» be able to provide a clear data audit trail
» have a robust review process that identifies what went wrong and prevents a recurrence
» start a crisis management process that helps you talk to your stakeholders

Thinking through step by step what you would do if there were a breach will help ensure you have considered all the important factors.

The types of things you need to think about include such as:

» What immediate measure would I need to undertake to prevent further breaches e.g. do I need to remove my data from the web and if so how will I go about doing this?
» Which external organisations and individuals would I need to notify and how should I do this?
» How would I interact with the person or organisation who is involved in/created the breach?
» Are all my data handling, data sharing and release policies and procedures robust enough?

Topic 10 Summary

1. Develop a breaches policy in which you outline for example what a breach is and who is responsible.
2. Think through step by step what you would do if there were a breach, which will help you ensure you have considered all the important factors.
3. Anticipating the rare scenario of a data breach will give you some control over its impact.

Well done! You completed the course.

You have seen all 10 components of the Anonymisation Decision-making Framework. Find more information at the UK Anonymisation Network's website. You can also stay in touch with us by following UKAN on Twitter.

Please share the course with anyone who you think may find it useful. The link is http://ukanon.net/course

Important reminder

The course should have given you a taster of the types of things that you will need to think about if you want to anonymise data effectively. It should enable you to work with experts more effectively and will reduce the costs of employing expertise. We recommend that you seek expert advice before releasing/sharing data derived from personal data.