Document Title	Code of Practice 12: Anonymisation of Data
Version	11/09/19
Author	Information Compliance Manager, VCO
Owning Department	Vice-Chancellor’s Office
Approval Date	26/09/19
Review Date	26/09/21
Approving Body	Information Assurance and Security Committee
Relevant to	All academic and professional services staff

All information which University of Greenwich staff holds and processes should be processed according to Data Protection and Freedom of Information laws.

1. What is anonymisation?

Anonymisation is the process of turning personal data into a form which does not identify individuals and where identification is no longer able to take place.

However, you should exercise caution when attempting to anonymise personal data. Organisations frequently refer to personal data sets as having been ‘anonymised’ when, in fact, this is not the case. You should therefore ensure that any treatments or approaches you take truly anonymise personal data. There is a clear risk that you may disregard the principles of the General Data Protection Regulation (GDPR) in the mistaken belief that you are not processing personal data.

In order to be truly anonymised under the GDPR, you must strip personal data of sufficient elements that mean the individual can no longer be identified. However, if you could at any point use any reasonably available means to re-identify the individuals to which the data refers, that data will not have been effectively anonymised but will have merely been pseudonymised. This means that despite your attempt at anonymisation you will continue to be processing personal data. Particular care should be taken in situations where individuals could be identifiable through combinations of variables in the data or through the data being linked to variables in other publicly available data sources, even if identifiers are removed. In such cases, the data will be pseudonymised rather than anonymised.

2. Some methods of anonymisation

a) Redaction – this removes certain pieces of information from data or documents, for example individuals’ names could be removed or “blacked out” from documents.

b) Reduction – similarly, this removes certain elements of the data, or reduces it to a level whereby fewer data are shown such that individuals cannot be identified from it.

c) Aggregation – in aggregation, data is displayed as totals or percentages only, so that no data relating to or identifying any individual is shown. Small numbers in totals are often suppressed through ‘blurring’ or by being omitted altogether.

d) Barnardisation – this refers to a method of disguising the data – in statistical information this could be produced by randomly adding or subtracting 1, for instance.

e) Blurring – of, for instance, video footage to disguise faces.

f) Electronically disguising, or re-recording audio material, for instance with an actor's voice.

g) Replacing – for example, changing the details in a report e.g. of place names, dates etc. Peoples' names could be replaced by pseudonyms. However, "pseudonymised" data are considered to be personal data under the GDPR – see more information below.

3. Pseudonymisation

Pseudonymisation is the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Pseudonymisation typically involves replacing names or other identifiers which are easily attributed to individuals with, for example, a reference number. Whilst you can tie that reference number back to the individual if you have access to the relevant information, you put technical and organisational measures in place to ensure that this additional information is held separately. Pseudonymisation does not remove all identifying information from the data but merely reduces the linkability of a dataset with the original identity of an individual. Pseudonymous data still allows for some form of re-identification (even indirect and remote), while anonymous data cannot be re-identified because all potentially identifying variables are removed.

Pseudonymising personal data can reduce the risks to the data subjects and help us to meet our data protection obligations. However, pseudonymisation is effectively only a security measure. It does not change the status of the data as personal data. Recital 26 of the GDPR makes it clear that pseudonymised personal data remains personal data and therefore within the scope of the GDPR. Pseudonymising data is useful for complying with data minimisation and security, but individuals may still be able to be identified from it.

Even though pseudonymised data does not identify an individual, in the hands of those who do not have access to the 'key' to the data, the possibility of linking several anonymised datasets to the same individual can be a precursor to identification, so every care must be taken.

4. Research data

Pseudonymisation is a useful tool to use when processing research data. Ways of pseudonymising research data could be:

- Separate the research data from the identifying details of the participant
- Give the data a code and attach the code to the identifying details of the participant
- Allow participants to choose codes / passwords so that they could be allowed access to their data if necessary / withdraw from the project within certain timescales

This is not anonymisation of data, it is pseudonymisation of data. Ensure that you only publish the data in truly anonymised form. We need to ensure that the data never causes damage or distress to individuals.

5. Data or information being used for training purposes

Never use data or information about real people who can be identified from it for training purposes. That could be in situations where you are demonstrating by, for example, a Powerpoint presentation or other on-screen or online demonstration in a classroom, or one-to-one, or in training materials which could include guidance documents, videos or suchlike. Always anonymise your demonstration data if you are basing it on real data – refer to the points above on how to make sure you are doing this correctly.

For further advice contact compliance@gre.ac.uk.