

How Functional Anonymisation can help Interpret & apply the GDPR 2016/679 Model of Personal Data



Dr Elaine Mackey & Dr Karen Mc Cullagh

Presentation Outline

□ **GDPR MODEL OF PERSONAL DATA**

- What does the model look like?
- The case of Pseudonymisation

□ **FUNCTIONAL ANONYMISATION**

- An approach: it can help one to determine the status of data as personal or non personal
- Case example: applying Functional Anonymisation to a data situation

General Data Protection Regulation 2016/679

□ On 25th May 2018, the GDPR 2016/679 will repeal and replace Directive 95/46/ec

Key features:

- Greater Accountability
- Stronger sanctions
- Enhanced data subjects rights

Directive 95/46/ec



Personal data

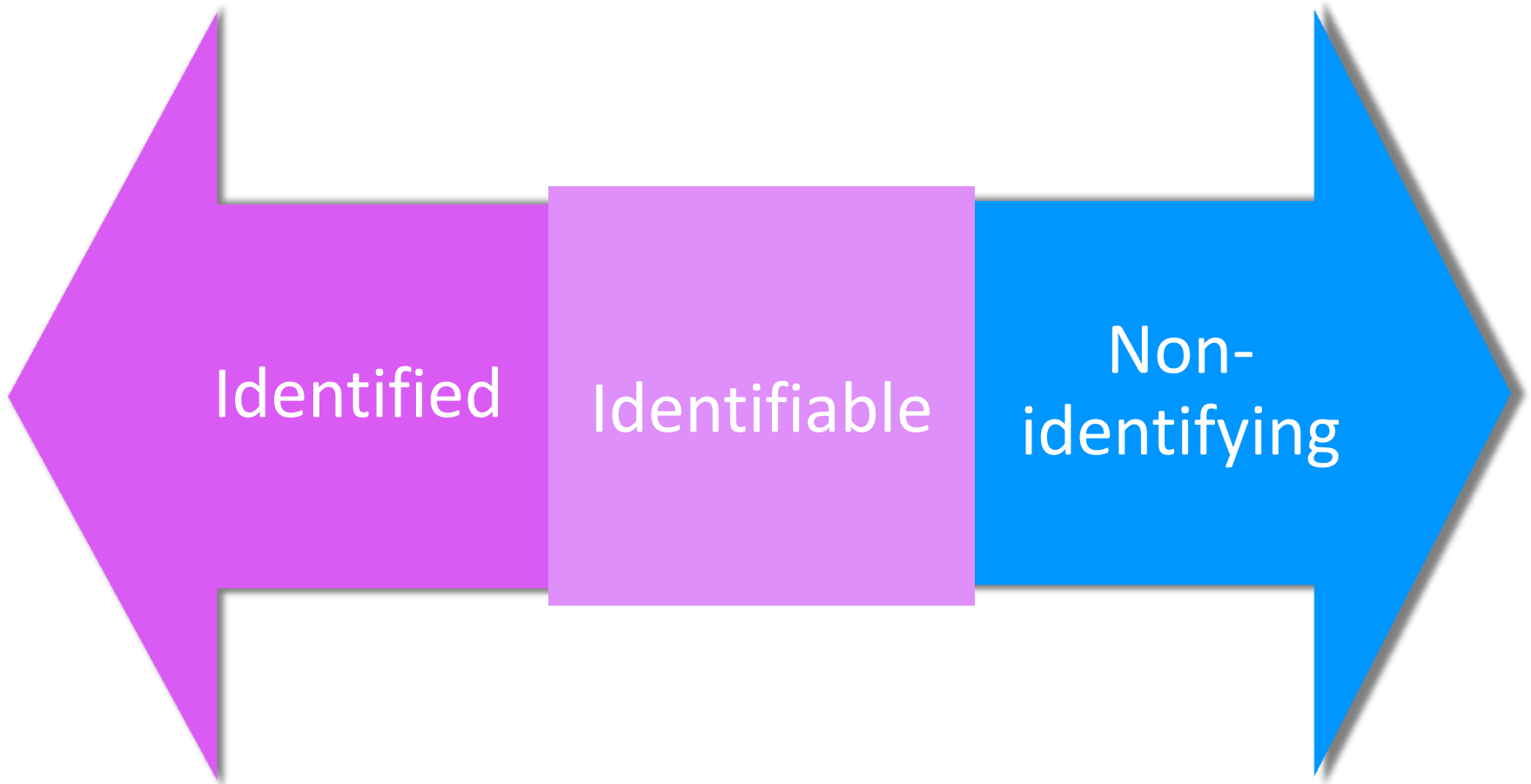
*'any information relating to an **identified** or **identifiable** natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity'* Art 2(a)



Anonymous data

'the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable' Recital (26)

Identifiability Continuum



GDPR Model of Personal Data

- ❑ Effect of including a new term:
pseudonymisation?
- ❑ Does it represent a change in the personal data model?

For example:

- Does it establish a tripartite model?
- Does it represent no change to the bipartite model?

GDPR 2016/679



Personal data

*any information relating to an **identified** or **identifiable** natural person; ... who can be identified, directly or indirectly, ... by reference to an identifier such as **a 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'* (Art 4(1))



Anonymous information

*'The principles of data protection should therefore not apply to **anonymous information**, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable* (Recital 26)

Definition Art 4 (5)

*'Pseudonymisation' means 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'*

Recital 26, pseudo=personal

*‘Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an **identifiable natural person**’*

Recital 28, reduced risks

*'The application of pseudonymisation to personal data can **reduce the risks to the data subjects** concerned and help controllers and processors to meet their data-protection obligations'*

'The explicit introduction of 'pseudonymisation' in this Regulation is not intended to preclude any other measures of data protection'

Recital 29, incentive

*'In order to create incentives to apply pseudonymisation when processing personal data, measures of pseudonymisation should, whilst allowing **general analysis**, be possible within the same controller when that controller has taken technical and organisational measures necessary to ensure, for the processing concerned, that this Regulation is implemented, and that additional information for attributing the personal data to a specific data subject is kept separately'*

'The controller processing the personal data should indicate the authorised persons within the same controller'

Incentives: relaxed data controller obligations

Pseudonymisation may


1. *facilitate the processing of personal data beyond original collection purposes. (Art 5, Art 6(4), Art 6 (4)(e))*
2. *provide a 'safeguard' for processing personal data for scientific, historical and statistical purposes.(Art 89(1), Art 25(1))*
3. *exempt a data controller from data subject access, rectification, erasure or data portability rights, provided that the Pseudonymisation prevents a controller from identifying a data subject (Art 11). Note: If a data subject provides the controller with additional information that allows them to be identified in the data set, they must be permitted to exercise those rights*
4. *help a data controller meet data security requirements (Art 32(1)(a))*
5. *reassure data subjects that a data controller has implemented appropriate safeguards "both at the time of the determination of the means for processing and at the time of the processing itself." (Art 25(1)) i.e. data protection by design, rather than data protection as an afterthought*

GDPR Model: any changes from the Directive?

- ❑ A more complex bipartite model
 - Pseudonymisation is **NOT** a new category of data
- ❑ Pseudonymisation is as a process
 - Reflects position in some national laws introduced to give effect to Directive 95/46/ec e.g. Germany Bundesdatenschutzgesetz (BDSG) [*Federal Data Protection Act*], Section 3a. (data minimisation)

Key Points

- ❑ At the **definition level**, *identified* data and *identifiable* data are considered **equivalent** i.e. Personal data
- ❑ At the **practical level**, *identified data* and *identifiable* data **are not** treated as **equivalent**. – i.e. some compliance obligations are relaxed for data that has undergone pseudonymisation

Functional Anonymisation: an 
approach to help determine
whether data are personal or non
personal

What you need to know

1. Identifiability
2. Identifiability and the issue of re-identification risk
3. The test for determining whether an individual is identifiable
4. Perspectives on determining identifiability & assessing re-identification risk



**Identified
and
Identifiable
Personal Data**



**Non
Identifying
Anonymous
Information**

**Continuum of Identifiability
Re-identification Risk**

Identifiability & Risk

There are 2 positions on the issue of risk

1. Absolute Anonymisation

- Zero risk of re-identification

2. Risk Based Approach

- Zero risk not possible
- Most widely accepted approach
- Data controller should ensure that the risk of re-identification is remote

Determining whether a person is identifiable

*‘To determine whether a natural person is identifiable, account should be taken of all **the means reasonably likely to be used**, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly’*

*‘To ascertain whether means are reasonably likely to be used to identify the natural person, **account should be taken of all objective factors**, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments’ (Recital 26)*

Determining Identifiability

Determining identifiability is complex and shaped by the particular approach you take

Example 1:

John Smith, 4 Willlowbrook, BL1 5CV, DOB: 1/1/79, male, diabetic, treatment episode on 5/1/18, Bolton General Hospital

- ❑ Relatively easy to assign a status to example 1: 'Directly identifying personal data'

Determining Identifiability & Re-identification Risk

□ Example 2:

39 year old, male, diabetic, treatment episode 5/1/18, Bolton General Hospital

□ Example 3:

35-45 year old male, diabetic, treatment episode January 2018, North West

For examples 3 & 4 can you determine identifiability by looking just at the data?

Don't just look at the data

- ❑ You cannot determine identifiability and re-identification risk by looking just at the data in front of you

Traditional Perspective

- ❑ Dominant approach
- ❑ Risk seen as originating from and largely contained within the data to be shared
- ❑ Looks first and foremost at the data to determine its identifiability and assess risk

Traditional Perspective

- ❑ Data context increasingly seen as important
- ❑ However, there is a tendency towards thinking the data environment is too difficult to gauge so focus remains on the data

Data Environment Perspective

- ❑ Recent approach
- ❑ To determine identifiability & assess re-identification risk one has to look at both **data environment** and **data**

Data Environment Perspective

- Re-identification risk arises from the interaction between: **The dataset,**
 - **People; Other data; Infrastructure & Governance**

Mackey, E. & Elliot, M. (2013) Understanding the Data Environment', *XRDS*, 20(1); 37-39

- This approach underpins Functional Anonymisation



Principle of Functional Anonymisation:
you can only determine whether data are
personal or not in relation to their environment

Functional Anonymisation

□ In practice what this means is:

**Look at the Data Environment
first**

**then look at the Data in
relation to that Environment**

Functional Anonymisation

Functional Anonymisation was first described by

- ❑ Dibben, C., Elliot, M., Gowans, H. and Lightfoot, D. (2015) 'The Data Linkage Environment' in Methodological Developments in Data Linkage, eds. Harron, K., Goldstein, H & Dibben, C. Wiley UK.

Later incorporated into

- ❑ Elliot, M., Mackey, E., O'Hara, K. and Tudor, C. (2016) 'The Anonymisation Decision-making Framework' <http://ukanon.net/ukan-resources/ukan-decision-making-framework/>

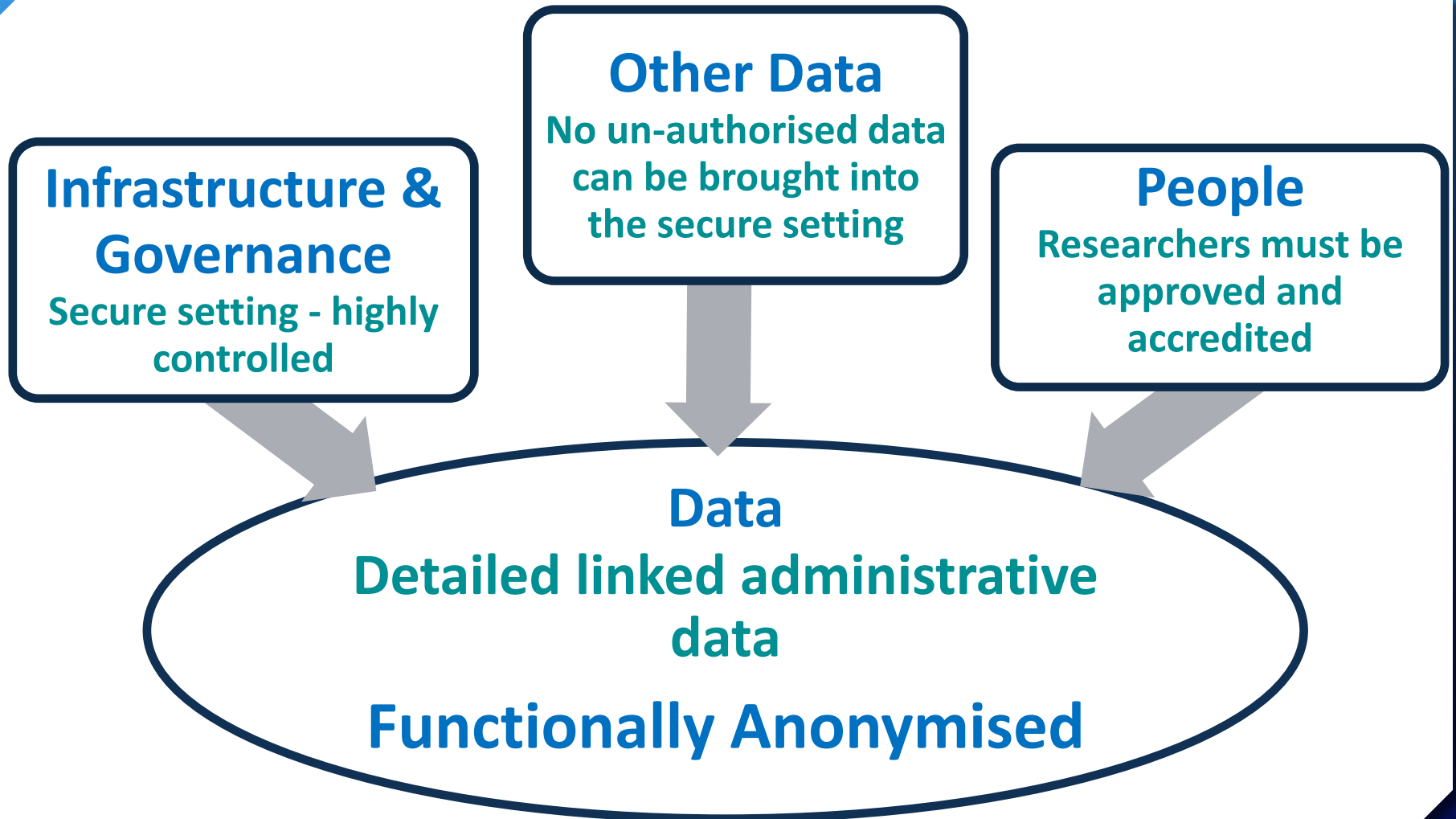
The Case of the ADRN

<https://www.adrn.ac.uk>



ADRN

Controls placed on data & environment





**Identified and
Identifiable
Personal Data**



**Non Identifying
Anonymous
Information**



**Re-identification
Risk Remote**

Concluding remarks

- ❑ Whilst fundamentally the model of personal data does not change under GDPR – it is more complex
- ❑ Functional Anonymisation is an approach that can help you navigate the complexity of GDPR model in terms of categorising data
- ❑ Look at the data environment first and foremost – then at the data in relation to its environment

Contact details

Dr Elaine Mackey, Research Associate,
University of Manchester,
elaine.mackey@manchester.ac.uk

Dr Karen Mc Cullagh, Lecturer in law,
University of East Anglia,
k.mccullagh@uea.ac.uk