CO-CONNECT & Data Protection

Safeguarding data

November 2021

Introduction

This document is to brief Data Partners on the measures that the CO-CONNECT Team will be taking to protect and safeguard data, referred to in CO-CONNECT Implementation Document, Data Governance and Security Controls. It will outline the measures that CO-CONNECT will be implementing to ensure that data remains secure and legally compliant, as well as providing additional information Data Partners may wish to include in the Data Protection Impact Assessment (DPIA) for the CO-CONNECT project.

Background

Under data protection legislation and the duty of confidentiality[1], the key determinant when processing 'personal data' or 'confidential patient information'[2] is whether individuals could be identified directly or indirectly from the processing of that data and/or information. Where individuals cannot be identified, the data/information is rendered anonymous and falls outside both the data protection legislation and the duty of confidentiality.

In practice, the MRC guidance[3] states:

> *"there is a continuum of identifiability from complete anonymity at one end, through jigsaw identifiability, and on to directly identifying data/information at the other. Although identifiability is a continuum, the law is binary; data/information is considered either identifiable or anonymous."*

In determining identifiability, recital 26 of the GDPR states that:

> "…account should be taken of all the means reasonably likely to be used…to identify the natural person directly or indirectly."

Thus, a key element in complying with the law, is accurately determining whether individuals are identifiable.

CO-CONNECT: Privacy by Design

The CO-CONNECT project has established protocols and measures to ensure that individuals are not identifiable because the data has been rendered anonymous throughout all processing. These include the use of

- One-way irreversible hashing algorithms applied to 'linking keys' or core identifiers such as NHS and CHI numbers. These are applied at source and prior to any interaction with CO-CONNECT software[4].
- Obfuscation. Particularly of potentially identifiable information such as date of birth and 'event' dates[5].
- Platform aggregation within the software provided by CO-CONNECT. This ensures that only aggregated data is presented to the user searching the system, but the ability to check for duplicates between datasets is retained.
- Low number suppressions. This ensures that when a result returns less than a threshold set by each Data Partner, it will return no results.

[1] Principally the General Data Protection Regulation (GDPR) and Data Protection Act 2018.
[2] These are the terms used in the data protection legislation, and NHS Act 2006 which defines confidential patient information.
[3] https://mrc.ukri.org/documents/pdf/gdpr-guidance-note-5-identifiability-anonymisation-and-pseudonymisation/
[4] For further information please refer to the document titled *CO-CONNECT and Data Anonymisation*.
[5] For further information please refer to the document titled *CO-CONNECT and Data Anonymisation*.

- Rounding of results. This ensures that it is not possible to find edge cases by asking multiple questions.
- Search criteria can only be constructed from pre-defined fields. This ensures that users can only query data that has been authorised.
- Drag and drop interface. This ensures that no queries are written by the user and query creation is strictly controlled.

Identifiability under the GDPR and the Duty of Confidentiality

Data Protection Legislation

Data protection legislation covers personal data. Personal data is defined in Article 4 (1) GDPR as

> *"…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 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;"*

Recital 26 of the GDPR further explains how identifiability is determined

> *"The principles of data protection should apply to any information concerning an identified or identifiable natural person. 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. 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. 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. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."*

Thus, in determining identifiability account should be taken of all the means reasonably likely to be used by the controller or by another person to identify the natural person directly or indirectly. The assessment of what is reasonably likely should be made on the basis of the particular circumstances of the data processing.

The Information Commissioner's Office (ICO) accepts that where data is transferred to organisation 2, but data that could permit identification stays with organisation 1, and there are technical and organisation measures in place such that organisation 2 does not have access to the identifying data, and therefore do not have the means reasonably likely to identify individuals, the data could be considered to be anonymous in the hands of organisation 2. However, the process of rendering the data anonymous for transfer to organisation 2 would be considered processing of personal data. As such, organisation 1 would need to comply with data protection legislation in undertaking the process of anonymisation in order to transfer anonymous data to organisation 2.

Duty of Confidentiality

The duty of confidentiality covers confidential patient information. Confidential patient information is defined in Section 251 (11) of the NHS Act 2006

For the purposes of this section, patient information is 'confidential patient information' where –

a) *The identity of the person in question is ascertainable –*
   i. *From that information, or*
   ii. *From that information and other information which is in the possession of, or is likely to come into the possession of, the person processing that information, and*
b) *That information was obtained or generated by a person who, in the circumstances, owed an obligation of confidence to that individual.*

Thus, the identifiability of the individual is the key to determining whether the information is protected by the duty of confidentiality. Where the identity of individuals cannot be ascertained, the information is rendered anonymous and the duty of confidentiality does not apply. The assessment of whether it is likely that the person processing the information will come into possession of information which would enable identification, should be made on the basis of the particular circumstances of the data processing.

Please note in Scotland Common law of confidentiality is used as set out in the NHS Code of Practice (Scotland) Version 2.0

CO-CONNECT & Identifiability

The protocols and measures put in place by CO-CONNECT ensure that data received from Data Partners cannot reasonably be identified by either the software it provides, or by the wider CO-CONNECT project team. As such, and in-line with the above guidance, the data that CO-CONNECT receives and subsequently processes will be considered anonymous and not subject to data protection and/or duty of confidentiality legislation.