Developers guidance

Proof-of-concept: using anonymous or artificial health data

Downloaded on October 18th, 2024

This is required guidance

It is legally required and it is an essential activity.

This Guide covers:

- England

From:

- Health Research Authority (HRA)

Last reviewed: 20 January 2023

Proof-of-concept: using anonymous or artificial health data

Reviewed by: Health and Care IG Panel

Often, developers will use anonymous or artificial data (that does not, or no longer, relates to identified or identifiable individuals) during their proof-of-concept stage. The data you choose should still be appropriate to the context of the technology you are investigating.

Benefits of using anonymous or artificial data during proof-of-concept:

- it can speed up your technology's route to market, because you will not need to wait for consent or approvals
- the results of your proof-of-concept can justify why you need to use personal data at a later stage
- it reduces the risk/detriment to patients if a data breach occurs while you are developing your technology

Data minimisation

Designing your technology to minimise processing of personal data to what is needed is also a legal requirement under UK GDPR. Even if personal data can be justified as necessary, it must be minimised to only the amount of data needed to carry out the project purpose at that stage and no more. Data minimisation is also a requirement of the Caldicott Principles relating to confidential data in health and social care, which you must follow. For example, the Caldicott Principles state that the minimum data possible should be used and accessed strictly on a 'need to know' basis. For more information, see step 7 in complying with the UK GDPR.

When developing a digital technology, therefore, you need to consider the kind of data that is the best fit for each stage of development. This is to make sure you only use what is necessary and proportionate to achieving your purposes at each stage.

The UK GDPR also requires you to put in place technical and organisational measures to fulfil the data protection principles and safeguard individual rights. This approach is described under Article 25(1) and (2) as 'data protection by design and by default'. Your first consideration should be: what kind of data do I need for the development stage of my technology? At proof-of-concept stage, using artificial or anonymous data may be a better choice than using personal data or confidential patient/service-user data and will reduce your level of risk, as described below. Building data protection functionality such as access controls, audit trails, and appropriate privacy notices for the intended users is an effective way of meeting this data protection legal requirement.

 Downloaded on 18/10/2024, for the latest information, check the website as printed documents may be outdated.

Get more information: Data protection by design and default | ICO

How to process anonymous or artificial data

Usually, you do not need consent or approval to process data that has been rendered anonymous (including through synthesis), or artificial data. This is because, when a person cannot be identified from data, its use is not subject to the UK common law duty of confidentiality or data protection legislation.

Anonymous data

Data rendered anonymous data is no longer considered personal data. See the ICO's guidance on what is personal data?

Important note: Determining whether the data you wish to use is personal data or not is your responsibility and you should carry out your assessment helped by the latest guidance provided on the ICO website. You should check the ICO website from time to time as new guidance becomes available.

In this process of anonymising personal data, an organisation has modified data in order to share it with a third-party organisation, while also putting in place additional processes and other appropriate safeguards to prevent the use of means reasonably likely to identify the individual in the third-party's hands, and thus make sure it meets effective anonymisation requirements. A key benefit of this is that it reduces risk of a data breach of personal data, which could cause harm to patients and service users.

A lawful basis is required to anonymise personal data (see step 4 in complying with the UK GDPR: Have a lawful basis for processing health data). In particular, a lawful basis under the common law duty of confidentiality is required to share confidential patient and service-user information with someone who would then apply anonymisation processes to the data. If the person receiving the data is not part of the direct care team looking after the individual, there will be a disclosure of confidential information (albeit solely for the purposes of anonymising it) and a legal basis to lift the common law duty of confidentiality is required. This would likely apply to a technology developer. In such circumstances, you must obtain the prior explicit consent from the individual, unless there is another legal basis available to you, as described further below.

Downloaded on 18/10/2024, for the latest information, check the website as printed documents may be outdated.

Important note: this type of consent (explicit consent from an individual to permit confidential information to be shared outside the team directly caring for them) is separate from UK GDPR consent. However, the rules on consent do not conflict. This is because they are about consent for different things under 2 different sets of regulations that were created to work together without tension. For more on this distinction, see the NHS Transformation Directorate's guidance on consent and confidential patient information and the HRA's guidance on consent in research.

Important note reminder: if you are reliant on using anonymous data to fall outside the law, you must make sure the data you want to use has been rendered anonymous, and you have evidence to demonstrate that it is no longer personal data, before using or disclosing it. Any onward transfer of (or remote access to) the data may change its status to be personal data again, depending on any additional information and means available to the onward recipient. Therefore, the effectiveness of your anonymisation strategy must be determined on a case-by-case basis, using the latest guidance provided on the ICO website.

A note on pseudonymisation

Pseudonymisation is a technique applied in circumstances when the link between individuals and the data that relates to them needs to be reduced but not removed entirely. It involves replacing information in a data set that directly identifies an individual. For example, it could involve replacing an NHS number, a name, or an address, with a unique number or code (a pseudonym). This has the effect that those receiving it cannot identify an individual directly from that data without access to additional information held separately and securely elsewhere (for example, the 'key' that would enable matching the pseudonym to the removed direct identifiers).

UK GDPR legislation applies to personal data. The UK GDPR considers that pseudonymisation is not an anonymisation technique. This is because, by itself, applying the technique does not render personal data anonymous in the hands of those receiving the information. More is required, such as putting in place a data-sharing contract and other appropriate safeguards to prevent reidentification by the recipient. The determining factor is whether or not the recipient can use the modified data (whether on its own, or in conjunction with other available data using reasonable means) to identify an individual.

If you are using pseudonymised data, make sure you understand your legal obligations described in how to process health data.

Get more information:

 Downloaded on 18/10/2024, for the latest information, check the website as printed documents may be outdated.

ICO guidance on anonymisation, pseudonymisation, and privacy enhancing technology

Downloaded on 18/10/2024, for the latest information, check the website as printed documents may be outdated.