European Data Protection Board

February 2025

Pseudonymisation

when and how to apply it

Pseudonymisation is a way for organisations **to protect personal data by making it more difficult to link it back to individuals. It can help organisations to meet some of their GDPR obligations regarding, for instance, security or data protection by design** (incorporating privacy measures from the start of data processing), **and by default** (processing data with the highest level of privacy protection).

It works **by replacing personal data with information that does not directly identify a specific individual.** To reconnect the data to a person, extra information is needed, like a lookup table or a secret key. As pseudonymisation is a technical and organisational measure to reduce the risks for individuals and meet data protection obligations, it does not need its own separate legal basis. The legal basis for the processing of the personal data extends to the pseudonymising transformation.

The EDPB guidelines on pseudonymisation help controllers choose the best methods to modify original data, safeguard pseudonymised data from unauthorised identification, and effectively manage user rights when handling pseudonymised data.

The guidelines clarify two key points:

1 **Are pseudonymised data still considered personal data?**

Pseudonymised data, which can be linked back to an individual using additional information, is still personal data. This remains true when another party keeps the additional information.

2 **How does pseudonymisation help organisations process data under the GDPR?**

Pseudonymisation has several advantages, including reducing both security risks and the risks of using data for a purpose that had not been agreed. It can also make it easier for organisations to use legitimate interest as a legal basis for processing data.

Pseudonymisation is not always required

Depending on the nature, scope, context, purposes, and risks of data processing, organisations may choose to use pseudonymisation to meet some EU data protection requirements. Pseudonymisation is particularly useful regarding the data **minimisation principle** (only collecting and processing the minimum necessary data), implementing data protection **by design** and ensuring data protection **by default**. It may also help to ensure security. In some cases, EU or national laws may mandate the use of pseudonymisation.

EXAMPLE 1

Context and purpose of processing

A company offers an app that provides medical advice based on user-reported symptoms. One division is responsible for quality control, using data collected with users' consent to check if the advice aligns with medical guidelines and to identify patients who need to be notified about incorrect advice.

Main objectives: maintaining the connection between data and individuals while reducing identifiability to ensure data minimisation, privacy by default, and confidentiality of the data.

One solution:[1] the quality control division processes only the information they need to ensure best practices were applied, ie only pseudonymised data received from the app's backend. Only in the rare case where an individual needs to be informed of something, the division shares the pseudonym and message with the team that has access to the full user's information. This team can then identify the user and send them a notification.

EXAMPLE 2

Context and purpose of processing

A data centre collects health data from hospitals about participants in a long-term research project, along with data on occupational exposure to health risk from a labour agency. The centre provides query results to studies upon approval and coordinates access to medical records for quality control. It also notifies participants of any significant risks identified during studies.

Main objectives: ensuring appropriate safeguards for a processing for scientific research purposes, collecting and linking data from different sources, maintaining a connection between institutions and individuals, while ensuring that data are not attributed to individuals by the data centre employees or research groups.

One solution:[1] a board ensures that research groups take steps to protect data and prevent the identification of individuals. Institutions request approval from the board to use the data for a specific research project. The data centre then provides pseudonymised data to research groups within a secure system based on their request. The trust centre holds the means to match pseudonyms. If raw data is needed for quality checks or to inform individuals about risks, the pseudonymisation is reversed at the trust centre, and the hospital that submitted the data will contact the individuals.

Read the guidelines

[1] This is a simplified version of one example of the Annex of the guidelines. Please refer to the full example from the Annex before taking a decision on a similar case for important implementation details and needed additional safeguard.