Anonymisation and anonymity of image and signal data in processing under the Act on the Secondary Use of Health and Social Data (522/2019)

The Expert Group on Data Protection in the Secondary Use of Health and Social Data (VN/23353/2022) has outlined the following on the anonymisation and anonymity of image and signal data when processing subject to the Act on the on the Secondary Use of Health and Social Data (552/2019, hereinafter the Secondary Use Act). This policy does not apply to genome and genetic data. The policy does not take a stand on the tools or technical solutions used for anonymisation.

Background

- The definitions used in the Secondary Use Act are firmly based on the definitions laid down in the EU's General Data Protection Regulation (2016/679). The definition of personal data in the GDPR, the interpretation guideline in recital 26 of the GDPR's introduction, the case-law of the European Court of Justice, the guidelines of the European Data Protection Board, the guidelines adopted by the Article 29 Data Protection Working Party (WP29) and approved by the European Data Protection Board, and the expert group's previous policy[1] on the definition of anonymous data are considered the starting points for anonymisation. The minimisation principle laid down in the General Data Protection Regulation (2016/679) must be taken into account, and the amount and accuracy of the data must be assessed in relation to the intended use. By limiting data and reducing the resolution of images, the risk associated with identification can also be reduced.
- According to the Secondary Use Act, the data permit authority is responsible for the anonymisation of personal data on the basis of a data permit, regardless of which party has granted the data permit. The data permit authority is responsible for the anonymisation of data to be disclosed and for disclosing it to the permit holder for processing, with the exception of the processing of data by statistical authorities in accordance with section 7 of the Secondary Use Act. The data permit holder may produce anonymous results to be published when the data permit authority grants the right to do so. However, the data permit authority is responsible for verifying the anonymity of the results to be published.
- The party processing the data must regularly monitor the risk related to identification from anonymised material. Data protection regulations must be applied to the processing of data if anonymous data becomes personal data.

Anonymisation and anonymity of image and signal data

- The anonymity of image and signal data should chiefly be examined from the recipient's perspective. From the perspective of the General Data Protection Regulation (2016/679), the data remains, in principle, personal data in relation to the original controller if the controller retains the original personal data. However, the same information may be anonymous to other parties[2]. When data is published, the group of recipients is unlimited. In this case, special attention should be paid to the adequate anonymisation of the data.

- In most cases it is in practice impossible to identify a person on the basis of an individual image or signal data pattern (e.g. an EKG without a timestamp), without searching for the image or signal data in the patient information system, which is as a rule prohibited. Such an image can be considered anonymous if the image does not contain the patient's identification data and the recipient cannot combine the image with the patient's identification data or other information to identify the person by means that could reasonably be used by someone.
- All identification data linked to an individual must be removed from anonymous image or signal data pattern. These include name, personal identity code, date of birth, exact age, date of exposure and other metadata fields that enable identification. The risk of identification can be reduced with limitations and the reduction of resolution.
    - General additional information (e.g., the patient's rough age group and classification data on the finding) can be attached to the anonymous image and signal data as long as the image and signal as a whole remain anonymous.
    - When assessing the likelihood of identification, special consideration should be given to the following types of information, which involve a significant risk of identification: facial image (as a rule to be considered recognisable), a magnetic image revealing the shape of a face, a dental chart, a fingerprint, the iris of an eye, a palm image, an image of an extremely rare finding and an image of a tattoo in the skin (this list is not comprehensive).
    - If the image or signal data pattern represents a feature that varies from measurement to measurement (e.g. blood pressure), the risk of identification is lower.
    - Anonymous image or signal data cannot be associated with pseudo IDs that the recipient can link to specific research subjects.
- According to the General Data Protection Regulation (2016/679) and the Court of Justice of the European Union's case-law, it should be assessed whether the party processing the data has legal means at its disposal that it could reasonably use to identify individuals. When assessing how reasonably easy it is to identify an individual, all objective factors, such as the costs of identification, the time needed for identification and the technology and technological development available at the time of processing should be taken into account. On the basis of the case-law of the Court of Justice[3], the expert group has adopted the view that unlawful means are not considered reasonable. Although, it is possible that the original controller could find e.g. exactly the same image or pattern in the image and signal data repositories in their possession, in which case linking to the identification data and thus identification is possible, the regulation on the processing of personal data will only apply if the controller starts processing such data which is processed for other purposes. In general, there is no legal basis for such processing by the controller or another party. Furthermore, the combination of data would in most cases also be unreasonably difficult from a technical point of view. An assessment of whether there is a legal basis for combining data should be made using case-specific discretion.
- When assessing the anonymity of image and signal data, case-specific discretion must always be used and an appropriate documented risk assessment must be performed as should a data protection impact assessment, where necessary. Case-specific circumstances related to anonymisation play a fundamental role in assessing whether a person can be identified and the effectiveness of anonymisation of personal data. The possibility of combining other data with image or signal data increases the risk of identification. The more information is attached, the greater the risk of individuals being identified. When assessing anonymity and its associated risks as a whole, other data sets containing the same image or signal data must be taken into account.

References

[1] Resolution: Pseudonymisation, anonymisation and the use of direct identifiers in accordance with the Secondary Use Act (522/2019). https://stm.fi/documents/1271139/2013549/Pseydonymisointi_anonymisointi_linjaus.pdf/a00821ea-0b18-ff44-ee0d-d7080215820a/Pseydonymisointi_anonymisointi_linjaus.pdf?t=1603108675223

[2]That the same data may be displayed to one actor as personal data while another actor does not have the reasonable means to combine the data with a particular person is based on the Court of Justice's case-law on the concept of personal data. The case-law contains an assessment on the means available to different parties for linking information to a certain person and the reasonableness of the means available to them. Based on the case-law, it appears necessary to assess the means available to the recipient. See the judgement of 19 October 2016 Breyer, C-582/14, EU:C:2016:779 by the Court of Justice, paragraphs 44-45 and the judgement of 9 November 2023 C-319/22, EU:C:2023:837 paragraph 49, which guide to assessing the situation from the perspective of various actors instead of the court outright stating how the personal data in question should be processed in all situations regardless of the means available to the parties processing the data. The expert group also notes that the ruling of the General Court of 26 April 2023, SRB v EDPS, T-557/20, EU:T:2023:219, paragraphs 101-105 (the judgement is not final when this policy is adopted) has come to the same conclusion.

[3]The expert group's interpretation was based on the case-law of the Court of Justice. See the judgement of 19 October 2016 Breyer, C-582/14, EU:C:2016:779 by the Court of Justice. According to paragraphs 45-46 of this ruling, it must be determined whether a dynamic IP address can be linked reasonably to the additional information held by the internet access provider in order to identify the data subject. This was not the case, for example, when the identification of the data subject is prohibited by law.