Comments in response to the EDPB's public consultation on the proposed guidelines on data anonymisation

Contributors: Winnie Dongbou Wamba, Lisa Fériol, Gauthier Chassang, Anne-Marie Duguet, Noémie Dubruel, Jean Herveg and Nathalie De Grove-Valdeyron.

Contribution on behalf of UNESCO Chair in Ethics, Science and Society Working Group "Ethics of digital technology, robotics and AI" and CERPOP, University of Toulouse, Inserm UMR 1295, BIOETHICS Team

General comments on the guidelines:

These guidelines are valuable in that they clarify and update the applicable requirements regarding anonymisation, taking into account all types of data without sector-specific or contextual qualifications.

However, these approaches appear to be particularly technical and raise several difficulties regarding their practical implementation, particularly in highly sensitive fields such as health research. Furthermore, whilst we understand the value of the broad scope sought by such a document, we emphasise the need to take into account particularly specific and sensitive data, such as genetic/genomic data, for which anonymisation is (virtually) impossible. In order to clarify the scope and substance of these guidelines, we believe that various key points need to be addressed. These include, in particular, situations involving the transfer of data outside the European Union; the definition of an 'insignificant' risk of re-identification; consideration of potential technological developments that may alter the possibilities for re-identification; and, finally, the fundamental consideration of respect for individuals' rights.

Several specific points requiring further examination have been identified:

1. **The issue of transfers of personal data outside the EU**.

Observations. Consistency with the EDPS's Recommendations 01/2020 on measures supplementing transfer mechanisms designed to ensure compliance with the EU's level of protection of personal data.

The two EDPB texts deal with the same situation in different ways.

Recommendations 01/2020. In Case 2 of Annex 2 on examples of supplementary measures (paragraph 85), the EDPS acknowledges that pseudonymisation constitutes an effective

supplementary measure where data transferred abroad cannot be attributed to an identified or identifiable individual, taking into account, in particular, the information that the public authorities of the recipient country may be deemed to hold and use. A pseudonymisation technique configured in such a strictly defined manner is regarded by the recommendations solely as an effective security measure enabling compliance with Chapter V of the GDPR on data transfers.

This results in an inconsistency in classification. The same factual finding – namely that no attribution is possible either by the data importer or by the authorities of the third country – leads, in Recommendations 01/2020, to pseudonymised personal data accompanied by an additional measure. In the draft Guidelines 02/2026 on anonymisation, it may lead to anonymous data for the data importer.

Furthermore, data transfer situations require consideration of the concepts of 'highly sensitive non-personal data' as defined in Article 5(13) of the Regulation (EU) 2022/868 on data governance, as well as the context of data sharing arising from the implementation of the European Health Data Space. This concept, as set out in Article 88 of Regulation (EU) 2025/327 on the European Health Data Space in the context of the transfer of non-personal electronic data to a third country, states that non-personal electronic health data made available to a health data user in a third country by the bodies responsible for data access could be considered 'highly sensitive' within the meaning of Article 5(13) of Regulation (EU) 2022/868 on data governance, 'where the transfer of such non-personal electronic data to third countries presents a risk of re-identification by means going beyond those reasonably likely to be used, in particular in view of the the limited number of natural persons to whom those data relate, the fact that they are geographically scattered the technological developments expected in the near future'. Clarification within the guidelines on how to assess this risk of re-identification would also be welcome in the context of the provisions set out in the EHDS.

Requests.

(a) **Relationship between the two texts.** Clarify the relationship between them, and specify whether Chapter V applies to the transfer where the data are anonymous from the perspective of the importer and the authorities in their country, but remain personal data for the exporter.

(b) **Clarification regarding the tests.** Clarify whether the criterion of 'all information that the authorities might have at their disposal and might use' and that of 'means that could reasonably be used' constitute a single test.

(c) **Method of assessment by public authorities.** Refer to, or include directly within, paragraph 30 of the draft guidelines on the assessment by public authorities, the examination of the third country's legislation and practice as provided for in Recommendations 01/2020, paragraphs 30 et seq.

(d) Clarify the scope of the concept of 'highly sensitive non-personal data' (what about anonymous data that might be highly sensitive?);

(e) The concept of 'highly sensitive' non-personal data introduced by Regulation (EU) 2022/868 on data governance in Article 5(13), in order to provide for specific limitations or conditions in the context of transfers to third countries that may entail 'a risk of re-identification of anonymised non-personal data', should be addressed within the framework of these guidelines. This would clarify how this risk should be assessed, and by whom, as the concept of controller liability cannot apply given that personal data is not involved.

(f) In relation to the concept of highly sensitive non-personal data, what are the implications for the definitions? This concept appears to suggest that a risk of re-identification may persist even though the concept of anonymisation is defined by the irreversible nature of de-identification. Clarification is required regarding the various concepts used in different texts: anonymised data, anonymous data, non-personal data? These terms could also be defined in the glossary.

(g) As the EHDS is based on the principle of access to anonymised data (Article 66(2) of the EHDS Regulation), unless the user demonstrates that the purpose of the processing cannot be achieved using anonymised data, could the specific issues relating to health data be addressed within the guidelines, at the very least through examples? The application of the criteria set out in the guidelines to specific cases involving data considered particularly sensitive under the EHDS Regulation (Article 51(4)), such as human genetic, epigenomic and genomic data, and other human molecular data (omics data), could also be clarified.

2. **The classification of public authorities and intelligence services as relevant entities to be taken into account when analysing the reasonable means that may be used to identify an individual (paragraphs 29 to 33)**

Comments. Paragraph 30 lists the following among the relevant entities:

- law enforcement authorities and national intelligence services, by virtue of their legal powers, financial resources and technical capabilities;
- foreign intelligence services, which may, moreover, not be subject to the same laws.

The final sentence of paragraph 29, read in conjunction with footnote 23, adds that, where the law provides a means of accessing additional information, such access objectively constitutes a means that is reasonably likely to be used.

This gives rise to three difficulties.

Absence of a relevance criterion. In the absence of such a criterion, a prudent data controller should include these parties in any assessment. Yet, by definition, these are the entities that hold the most ancillary information and possess the greatest technical capabilities. Including them by default reduces the contextual approach to a quasi-absolute test, which is difficult to reconcile with the approach established by the *EDPS v SRB* judgment.

The concept of a lawful means of access. The *Breyer* judgment (C-582/14, paragraphs 47 to 49) concerned a specific lawful means: a specific entity could, in specific circumstances, obtain specific information through the competent authority. A general power of access during investigations does not possess these characteristics.

Requests.

(a) **Relevance test.** A public authority or an intelligence service should only be considered a relevant entity if objective and documented evidence establishes a concrete possibility of access to the data in question. Examples include: a legal competence covering that category of data; a documented practice of access; or the location of the data or its provider within the relevant jurisdiction.

(b) **Concrete legal basis.** Specify, in paragraph 29, that only a concrete legal basis for access within the meaning of the *Breyer* judgment constitutes a means reasonably likely to be used. General powers that do not apply to the data in question do not constitute such a basis.

(c) **Presumption of lawfulness.** Confirm that, for the authorities of the Union and of the Member States subject to the rule of law, the presumption of compliance with the law set out in paragraph 32 applies, unless there is evidence to the contrary within the meaning of paragraph 33.

3. **The definition of the threshold for an 'insignificant' risk of re-identification and the methods of demonstration (paragraphs 8, 22, 26 and 87)**

Comments. The draft bases anonymity on a probability standard:

- **paragraphs 8 and 22:** the probability of identification need not be zero, but must be negligible in practice, in accordance with the judgment *in EDPS v SRB* (C-413/23 P, para. 82).
- **paragraph 26:** means are not reasonably likely to be used where identification is impossible, requires a disproportionate amount of time, cost or labour, or is prohibited by law (*SRB*, paragraphs 82 and 83; *Breyer*, C-582/14, paragraph 46).
- **paragraph 87:** a re-identification technique is effective if it produces a result that is sufficiently close and sufficiently reliable to distinguish the individual. It may be effective even if it succeeds in identifying only a single person.

However, the draft provides no guidance on how to demonstrate that a risk is negligible.

This gives rise to three difficulties.

Impossible to demonstrate. The 'negligible' standard is a legal criterion established by the Court, but the draft does not translate it into operational terms. Yet the data controller must document the anonymisation and its effectiveness (paragraph 41) and be accountable for compliance (Article 5(2), GDPR). Without a recognised method, the data controller can neither demonstrate nor contest that a risk is negligible.

No mention of quantitative approaches. Clinical research relies heavily on quantitative measures of the risk of re-identification. This is the case with the EMA's external guidelines on Policy 0070, which set a threshold of 0.09 [to be verified]. The draft does not state whether these thresholds constitute admissible evidence, nor under what conditions.

Unspecified unit of measurement. Paragraph 36 requires that the risk be negligible for each individual, whilst paragraph 87 acknowledges that a successful attack against a single individual is sufficient. The draft does not specify whether the risk should be measured on a record-by-record basis (maximum risk) or at the dataset level (average risk). This choice significantly alters the outcome for cohorts containing outliers.

Requests.

(a) **Status of quantitative measures.** Clarify that the insignificance of the risk is a legal standard, which no single figure can fully capture. Quantitative measures of risk should nevertheless be accepted as evidence in the documentation, provided that the method is justified and the context is taken into account.

(b) **Sector-specific thresholds.** Express a clear position on their compatibility with the guidelines, in particular EMA Policy 0070, and on the conditions for their use (contextual approach, controlled access). Furthermore, it appears necessary to take into account specific situations such as health research, for example (an area where there is a particular need to differentiate the anonymisation analysis according to the types of research, the categories of data processed, the purposes pursued and the parties involved, particularly in the context of complex research projects).

(c) **Unit of measurement.** Specify whether the risk is assessed per record or at the level of the dataset, in line with paragraph 36.

(d) **'Sufficient confidence'.** Clarify this concept within the meaning of paragraph 87, elaborating on Example 20 and Note 44, according to which an inference is relevant only if it significantly alters the prior belief. This comparative criterion could serve as a common basis for demonstrating that the risk is negligible.

4. **Technological developments and reassessment of re-identification risks (in particular paragraphs 13, 35 and 88)**:

Observations. Within the framework of the contextual approach and paragraph 35, the text calls for consideration of technological advances and the availability of additional data that may be useful for re-identification, and states that "if the likelihood of identification increases to a level that is no longer insignificant, the previously anonymous data should again be considered personal data and the controller will be accountable for any processing of that personal data". This raises several questions: is there a specific scale for assessing the level of re-identification risk to evaluate these developments? It would be useful to clarify that the 'controller' referred to in this passage should not be the original controller of the personal data but the one who received the anonymised data. Contractual arrangements should specifically provide for this scenario. As this party becomes a data controller subject to the GDPR, they must then be able to respect individuals' rights, and first and foremost their right to information. But will they have the means to do so, given that Article 11 of the GDPR does not require the processing of additional data solely for the purpose of complying with the GDPR?

Questions.

(a) (paragraph 88) Place particular emphasis on the impact of technological developments – both scientific and digital – on the potential for re-identification (this is particularly relevant in the case of data relating to rare diseases, for example). In paragraph 13: Clarify the importance of taking into account the anonymisation or potential re-identification techniques used, as well as the evolving nature of these techniques

(b) (paragraph 35) Clarify the obligations of data controllers in relation to the EDPB guidelines and the GDPR. In addition to including such clarifications within the guidelines, it might also be recommended that these be set out in a contract.

Similarly, the newly appointed controller must be able to clearly identify the options available to them. In our view, the new controller should refrain from sharing such data if they are unable to implement supplementary anonymisation safeguards in accordance with the guidelines; or, delete the data. The best practice of regularly reassessing the anonymity of data in light of these developments seems difficult to apply, as anonymised data is no longer covered by the GDPR and this would require potentially significant resources.

6. Rights of data subjects

Observations. We were surprised to note that the rights of data subjects are addressed rather late in the guidelines and appear to be considered in a piecemeal manner, even though they remain at the heart of considerations relating to data anonymisation, in line with the very essence of the GDPR.

Requests.

- (a) The reminder of data subjects' rights under the GDPR should be placed at the beginning of the document, linking the guidelines to the existing obligations of the various parties involved in the anonymisation of personal data (in particular, linking them to the right to be informed about the possible anonymisation of personal data – as this process constitutes the processing of personal data – as well as the right to object).
- (b) The guidelines could clarify the relationship between respecting individuals' rights and the concept of anonymisation, particularly where certain parties hold anonymised data in a specific context of use.
- (c) Emphasising the importance of individuals' rights appears particularly essential in sensitive areas such as healthcare. Example 4 of the guidelines cites the use of data from electronic health records shared between healthcare organisations – in such particularly sensitive situations, it is important to highlight the relationship with individual rights, as well as the sensitivity of the data in question, which must be taken into account when assessing the risk associated with the processing.

  Indeed, it is essential to recognise and ensure respect for individuals' rights, even when anonymisation is involved, because, on the one hand, anonymisation remains a form of data processing within the meaning of the GDPR and, on the other hand, to enable data subjects to retain the right to access and control the data concerning them, and ultimately to enable them to exercise their right to take collective action as foreseen in Article 80 of the GDPR.

  (d) The EDPB's guidelines should serve as a reference document for ensuring respect for individuals' rights, emphasising in particular the importance of informing data subjects where possible, especially in situations involving data sharing where anonymisation depends on the recipients and the context, which may evolve in line with technological advances.