Sonae

30th October 2026

EDPB Guidelines on Anonymisation

Sonae welcomes the opportunity to contribute to the public consultation launched by the European Data Protection Board (EDPB) on its draft Guidelines 02/2026 on Anonymisation. The EDPB's effort to update Article 29 Working Party Opinion 05/2014 in light of recent case law of the Court of Justice, notably EDPS v SRB (C-413/23 P), and of the significant technological developments of the past decade is a much needed and well received initiative.

From Sonae's perspective, anonymisation should not be seen merely as a technical exercise or as a way of placing data outside the scope of the GDPR. It is, above all, an instrument of data protection in its own right. It allows organisations to use data for analytics, innovation and research while reducing the risks for individuals, in line with the principles of storage limitation and data protection by design laid down in Articles 5(1)(e) and 25 of the GDPR.

Sonae therefore welcomes the relative and risk-based approach set out in the Guidelines, under which anonymity is assessed from the perspective of each relevant entity and the likelihood of identification need not be zero, but insignificant. This approach reflects the case law of the Court and provides a sound basis for a practical framework.

However, several provisions of the Guidelines, read together, risk undermining that premise. Combined with the broad range of entities, means and future technologies to be considered, they may lead to the conclusion that virtually no data can be regarded as anonymous in practice. This would discourage the use of anonymisation precisely where it is most valuable, as well as in the development and use of artificial intelligence.

The Guidelines also come at a time when the co-legislators are discussing, under the Digital Omnibus, targeted amendments to the GDPR, including to the definition of personal data. Coherence between both processes will be essential for legal certainty.

Against this background, Sonae's position is structured around the following key messages:

1. Anonymisation must remain an achievable and useful outcome

The Guidelines rightly recognise that the likelihood of identification must be insignificant in reality, not zero, and that means which are impossible, involve disproportionate effort or are prohibited by law are not means reasonably likely to be used.

This premise is, however, weakened by other provisions. Paragraph 21 considers data to be personal whenever it "cannot be ruled out" that a receiving entity has the means to identify individuals. Paragraph 29(e) refers to the technology available and to reasonably foreseeable developments, paragraph 30 includes foreign intelligence agencies, cybercriminals and unethical companies among the relevant entities, paragraph 31 discourages reliance on lack of motivation, and paragraph 92 anticipates that AI will further reduce the cost of re-identification.

Taken together, these provisions may be read as a presumption that data is personal until re-identification is shown to be impossible. Such a reading would, in effect, return to an absolute standard that the Court has not followed, and would be difficult to reconcile with the premise set out at the beginning of the Guidelines.

The consequence would be counterproductive. If anonymisation is perceived as legally unattainable, organisations will have fewer incentives to invest in it, and data that could be safely anonymised will continue to be processed in identifiable form.

This does not mean lowering the level of protection. A demanding but attainable standard protects individuals more effectively than one that is theoretically rigorous but rarely applicable in practice.

Therefore, the Guidelines should:

- **Confirm that the application of paragraph 29 does not amount to a presumption that data is personal unless re-identification is shown to be impossible;**
- **Align the test in paragraph 21 with paragraph 22, by referring to a likelihood of identification that is not insignificant, rather than to the mere fact that identification "cannot be ruled out";**
- **Clarify the threshold at which the likelihood of identification is considered insignificant, including examples where anonymity is maintained despite a remote possibility of re-identification by actors with extraordinary capabilities;**
- **Confirm that anonymisation remains an achievable and legitimate outcome under the GDPR.**

2. The assessment of means reasonably likely to be used should be proportionate and grounded in reality

Paragraph 29 identifies relevant objective factors, such as the properties of the data, the context of processing, access restrictions, the availability of additional information, cost and time, and available technology. However, the Guidelines give limited guidance on how these factors should be weighed together, particularly where some point to a possibility of re-identification and others indicate that it would be unlikely or would require significant effort.

Practical examples would be particularly valuable in this respect, notably for borderline cases that organisations frequently encounter, such as small populations, outliers, high-dimensional or highly granular record-level data and the possibility of combination with external sources.

The reference to reasonably foreseeable technological developments is particularly sensitive. Since technological potential is open-ended, this criterion may be read as encompassing any hypothetical future capability. It would be more effective to anchor it in concrete and demonstrable developments, combined with the periodic reassessment that the Guidelines already recommend as good practice.

Sonae welcomes the recognition, in paragraph 31, that the potential value of re-identified data should be taken into account where it offsets the costs and effort involved. By the same reasoning, where re-identification would have no meaningful value for any relevant entity, this should weigh in favour of an insignificant likelihood of identification, even if motivation alone cannot be decisive.

Therefore, the Guidelines should:

- **Provide guidance on how the factors listed in paragraph 29 should be weighed together;**
- **Include examples of borderline cases, notably small populations, outliers, high-dimensional data and combination with external sources;**
- **Anchor the reference to future technological developments in concrete and demonstrable developments, combined with periodic reassessment;**
- **Clarify that the absence of any meaningful value of re-identified data for the relevant entities weighs in favour of an insignificant likelihood of identification.**

3. Technical, organisational and contractual safeguards should be given clearer weight

In practice, anonymisation rarely relies on a single technique. It results from a combination of technical measures, such as aggregation, generalisation or perturbation, with access controls, segregation of datasets and teams, contractual commitments and legal prohibitions.

The Guidelines rightly call for limits to access to be assessed holistically, in paragraph 90, and recognise that, where data is subject to appropriate security controls, access constitutes a means reasonably likely to be used only for authorised users (paragraph 89). However, they give little indication of how much weight contractual and organisational measures can carry. Paragraph 34 requires contractual measures to be reliable, verifiable and enforceable, while noting that they may be revised or disregarded by the parties.

The argument that contracts alone cannot make data anonymous is understandable. However, where they are enforceable, auditable and supported by technical controls, contractual commitments materially reduce the likelihood of re-identification. This is particularly relevant for data-sharing arrangements with suppliers, partners and clients, such as the sharing of aggregated insights or collaboration in controlled data environments.

Without further guidance, there is a risk that such measures are given little weight in practice, even when combined with robust technical controls and effective oversight.

Therefore, the Guidelines should:

- **Provide guidance on the weight that reliable, verifiable and enforceable contractual measures can carry when combined with technical and organisational controls;**
- **Include examples of data-sharing arrangements, such as controlled data environments, in which combined measures support a conclusion of anonymity.**

4. Anonymisation within corporate groups and data ecosystems requires clearer guidance

Many organisations operate as groups of companies, and data increasingly circulates between several entities, each with its own role. The concept of a "chain of means", in paragraph 28, and the perspective of receiving entities are central in these structures, but their practical application remains unclear.

First, where several entities anonymise data at source and the anonymised data then circulates between them, it is unclear how the chain of means should be assessed. Where no entity in the chain has the means to re-identify, nor access to anyone who has them, and effective and verifiable segregation measures are in place, the data should be considered anonymous for all of them, even if re-identification would be theoretically possible for a third party with extraordinary resources or would require collusion for which there is no apparent reason.

Second, within a corporate group, Company X may receive an anonymised dataset from Company Y and later, in a different context, receive from Y a dataset containing personal data, without certainty as to whether the data subjects overlap. The Guidelines do not clarify whether the mere availability of the second dataset within the group is, in itself, a means reasonably likely to be used for the purposes of the No Linkage criterion, regardless of effective segregation between those handling each dataset. Nor do they indicate what likelihood of overlap would require the anonymity of the first dataset to be reassessed.

Finally, some systems are specifically designed to allow deterministic matching between datasets held by different entities without revealing the original identifiers. Guidance on how the anonymisation criteria apply to such systems would be valuable.

In all these situations, membership of the same group should not, in itself, determine the outcome. What should matter is effective access to additional information and the safeguards in place.

Therefore, the Guidelines should:

- **Clarify how the chain of means applies where no entity in the chain has the means to re-identify, nor access to anyone who has them, and the weight to be given to combined safeguards in that assessment;**
- **Confirm that the mere availability of a personal dataset within the same entity or group is not, in itself, a means reasonably likely to be used for the purposes of the No Linkage criterion, where effective technical and organisational segregation is in place;**
- **Clarify what likelihood of overlap between data subjects requires the reassessment of a previously anonymised dataset, and the relative weight of factors such as membership of the same group and the nature and scope of the data concerned;**
- **Provide guidance on how the anonymisation criteria apply to systems specifically designed to allow deterministic matching between datasets held by different entities, without revealing the original identifiers.**

5. Anonymisation should be treated as part of the original processing

Paragraph 38 qualifies anonymisation as a processing operation that requires a legal basis under Article 6 GDPR and, where applicable, a condition under Article 9(2), while presuming that these coincide with those of the preceding processing where anonymisation is part of the same activity and pursues the same purposes. This is a welcome step.

However, it does not go far enough. Treating anonymisation as an autonomous processing operation may require new assessments of lawfulness, purpose compatibility and information to data subjects. As a result, reducing risk would generate more compliance obligations than keeping data in identifiable form. This is at odds with the principles of storage limitation and data protection by design, and with the recognition of pseudonymisation as a technical and organisational measure.

Anonymisation at source should, therefore, be treated as the final stage of the original processing, or as a safeguard applied to it, without a separate assessment of lawfulness, compatibility or information. Compliance with the transparency requirements set out in paragraph 40 should be sufficient.

Where anonymisation serves a distinct subsequent purpose, for instance when data collected for operational purposes is later anonymised for analytics, innovation, research or product development, clear criteria are needed to distinguish between both situations.

Therefore, the Guidelines should:

- **Recognise that anonymisation at source may be treated as the final stage of the original processing, without a separate assessment of lawfulness, compatibility or information beyond paragraph 40;**
- **Explain the rationale for treating anonymisation as an autonomous processing operation, and the added value of requiring a separate assessment of lawfulness, purpose compatibility and information;**
- **Provide criteria to determine when anonymisation is part of the same activity and purpose and when it is linked to a distinct subsequent purpose, with examples involving analytics, innovation and product development.**

6. Documentation requirements should be proportionate and integrated into existing accountability tools

Paragraph 41 provides that controllers should document the anonymisation process, including the testing of supposedly anonymous datasets. Sonae agrees that adequate documentation is essential to demonstrate compliance.

However, the expected level of detail and technical evidence remains unclear. Without a reference framework, organisations may develop very different and, in some cases, excessively complex documentation processes, even for low-risk processing.

Any minimum framework should be indicative and proportionate to the risk and complexity of the processing, rather than a uniform checklist. It could cover elements such as the context and purpose of the processing, the characteristics of the source data, the technique applied, the relevant entities and their access, the additional information reasonably available, the assessment of the three criteria, the tests performed, any limitations identified and the residual risk.

Anonymisation assessments also overlap with existing accountability mechanisms, such as data protection impact assessments, records of processing activities, risk assessments and data protection by design documentation. Treating them as a separate instrument would duplicate information and fragment the documentation relating to the same processing.

Therefore, the Guidelines should:

- **Set out an indicative minimum set of elements for documenting anonymisation, proportionate to the risk and complexity of the processing;**
- **Clarify when formal technical testing or independent expert validation is necessary, and when a proportionate internal assessment is sufficient;**
- **Confirm that the demonstration of anonymisation may be integrated into existing DPIAs, records of processing activities and data protection by design documentation.**

7. Further guidance is needed on AI, synthetic data and legal transparency obligations

The relationship between anonymisation and AI remains one of the main areas of uncertainty for organisations. Article 10 of the AI Act requires training, validation and testing data used for high-risk AI systems to meet quality criteria and to be examined for possible biases. This may require a level of granularity that is difficult to reconcile with full anonymisation, and guidance on how both requirements can be met would be valuable.

Paragraph 82 recognises that inferences may be drawn from AI models and synthetic data. Further guidance on how the No Inference criterion applies to them, including the type of testing that would be considered sufficient, would help organisations adopt these techniques with greater confidence.

In addition, some disclosures are mandatory, such as information on pay or human resources indicators in sustainability reporting. Where they involve categories with few individuals, the No Inference criterion may conflict with legal transparency obligations. Guidance on how to reconcile both would support consistent compliance.

Therefore, the Guidelines should:

- **Clarify how anonymisation can be reconciled with the data quality and bias-detection requirements of Article 10 of the AI Act;**
- **Provide guidance on the application of the No Inference criterion to AI models and synthetic data, including appropriate testing methods;**
- **Address mandatory disclosures, such as pay or sustainability reporting, involving categories with few individuals.**

Conclusion

The draft Guidelines represent an important opportunity to provide organisations with a clear, modern and workable framework for anonymisation. Their relative and risk-based premise is the right starting point and reflects the case law of the Court of Justice.

From Sonae's perspective, the success of the Guidelines will depend on ensuring that this premise is carried through consistently. The assessment of means reasonably likely to be used should be proportionate and grounded in reality, combined safeguards should be given clear weight, and anonymisation should be treated as a risk-reduction measure rather than as a source of additional compliance obligations.

The Guidelines also come at a time when the co-legislators are discussing targeted amendments to the GDPR under the Digital Omnibus, including to the definition of personal data. Coherence with these ongoing legislative developments will be essential for legal certainty, and the Guidelines should be reviewed once that process is concluded.

Sonae's businesses span from food and specialised retail to financial services and technology, serving an ever growing number of customers in Portugal and abroad through physical stores, e-commerce, loyalty programmes and financial products. Across these activities, anonymisation is used to generate insights for operations, to collaborate with suppliers and partners, and to develop new services, including AI-based solutions, while protecting customers' privacy.

Given that, Sonae businesses are directly exposed to the questions raised by data circulating between several entities. Clear guidance on the chain of means, on intra-group data flows and on matching between datasets held by different entities would therefore be particularly valuable, both for Sonae and for many other European groups.

In this sense, the Guidelines should strengthen anonymisation as a practical tool that benefits both individuals and organisations. A demanding but attainable standard will encourage its use, reduce risk for individuals and support responsible innovation across the Europe.

Contacts:

Frederico Mota
Head of Public Affairs | Sonae
fpmota@sonae.pt | publicaffairs@sonae.pt

About Sonae:

Sonae is a Portuguese multinational and an active investment holding company, with a diversified portfolio of leading businesses in sectors such as retail, telecommunications, financial services, technology, property and new ingredients. Operating in around 60 countries, it is the largest private employer in Portugal and achieved a turnover of 11.4 billion euros in 2025. Sonae's commitment to sustainability has been recognised internationally, notably through its inclusion in the S&P Sustainability Yearbook 2026 and on the Carbon Disclosure Project (CDP), which distinguishes around 1% of the more than 23,000 companies assessed globally.

Sonae is registered in the European Union Transparency Register under number 687748032165-82.

Find out more at **www.sonae.pt**