Towards anonymisation based on context, actual risk and the information effectively disclosed

Leem's position in the context of the EDPB consultation on the draft Guidelines on Anonymisation

October 2026

PROJET

KEY MESSAGE | Leem calls on the European Data Protection Board to preserve a genuinely workable contextual approach, to establish the concept of *meaningful information* as a cross-cutting principle, and to fully recognise the combined role of technical, organisational and contractual safeguards. Anonymisation should be assessed in light of the means reasonably likely to be used and the actual information gain, rather than on the basis of hypothetical re-identification scenarios.

Leem welcomes the objective of the EDPB to clarify the conditions under which data may be considered anonymous. A protective and operational framework is essential to enable a responsible use of health data for research, innovation and improved patient care. However, several provisions in the draft Guidelines risk weakening the contextual approach and creating legal uncertainty for healthcare stakeholders.

1. Safeguard the contextual approach

The draft recognises that the anonymity of a dataset may be assessed differently depending on the stakeholders concerned, and that a contextualised demonstration may constitute an autonomous basis for establishing anonymity, alongside a "simplified" approach. This direction should be confirmed unambiguously, so as not to imply that one approach, namely the simplified approach, should systematically be preferred.

Leem recommends that the analysis be based on the means reasonably likely to be used by the relevant entities, their effective access and the information already available to them. General knowledge, the abstract existence of another data source, or future and indeterminate capabilities should not be sufficient to conclude that a person is identifiable.

2. Make "new and useful information" a cross-cutting principle

Leem welcomes the approach adopted in the draft Guidelines under the No Inference criterion, which is grounded in Bayesian information theory and the distinction between *prior belief* and

posterior belief. This approach rightly recognises that an inference should not undermine anonymity where it does not result in any significant information gain relating to an individual. However, this reasoning should be elevated into a broader, overarching principle applicable across all anonymisation criteria, including No Record Isolation and No Linkage.

Where a dataset provides no new knowledge regarding an individual's identity or attributes beyond the information already available to the relevant entity, there is no genuine information gain, no increased ability to single out or treat that individual differently, and consequently no impact on that individual's rights and freedoms. In such circumstances, the data should be regarded as anonymous within the meaning of Recital 26 GDPR.

This principle also addresses the issue of re-identification tautology, whereby an isolated record can only be associated with an individual because the entity already possesses all the information necessary to make that association. In such cases, the dataset itself discloses no new information and therefore should not be regarded as creating a risk to natural persons or as calling anonymity into question.

Proposed principle: zero information gain about a person = no personal information disclosed.

3. Avoid any mechanical application of population thresholds

The uniform application of a No Record Isolation threshold may prevent the disclosure of statistical results relating to small populations, even where no new information is disclosed. In exchanges between pharmaceutical companies, hospitals and pharmacies, a stakeholder may sometimes be able to link a record to a person solely because it already holds all the corresponding information.

Leem recommends that statistical thresholds remain factors to be considered, supplemented by an analysis of the context, the knowledge accessible to the entity concerned and the information gain actually produced.

4. Recognise organisational and contractual safeguards

Effective access restrictions, enforceable contractual obligations and robust organisational measures must be fully taken into account in the contextual analysis, alongside technical measures. Taken in isolation, they do not necessarily provide sufficient evidence of anonymity, but they directly contribute to determining whether identification is reasonably likely to occur.

This clarification is particularly necessary for research consortia, public-private partnerships and corporate groups. The existence of joint controllership, participation in a consortium or an equity relationship should not automatically preclude anonymity for an entity that does not have effective access to the data or information enabling identification.

5. Clarify the proposed approach to synthetic data and artificial intelligence

The No Inference criterion should be clarified for synthetic data and the training of artificial intelligence models. The concept of "*new information about a particular individual*" should

clearly distinguish between reconstructing information relating to a real person, producing new information through a model, and generating synthetic data that disclose no new information about a specific individual.

Otherwise, the Guidelines could undermine existing models and create disproportionate uncertainty for health research and innovation. The principle of "*new and useful*" or "*meaningful*" information should apply consistently to these processing operations.

6. Prioritise durable principles, supplemented by sector-specific use cases

Leem recommends consolidating general, technology-neutral and sufficiently durable principles in the Guidelines. Parallel work with the sectors concerned could usefully supplement them with examples relating to the dissemination of statistical reports, interactions with hospitals and pharmacies, research consortia, secondary use of health data, synthetic data and artificial intelligence.

Leem's priorities

1	Confirm that a rigorous and documented contextual analysis constitutes an autonomous route for demonstrating anonymity.
2	Integrate the concept of *new and meaningful information* into the three No Inference, No Record Isolation and No Linkage criteria.
3	Assess identifiability in light of effective access, reasonable means, and technical, organisational and contractual safeguards.
4	Clarify the application of the Guidelines to complex environments (consortia, corporate groups and public-private partnerships) and to the development of synthetic data and AI models.

Conclusion | A consistent approach based on context, actual risk and the information effectively disclosed will preserve a high level of protection for individuals while ensuring the scientific utility of data and legal certainty for stakeholders.

PROJET