Ecommerce Europe feedback on EDPB guidelines 02/2026 on anonymisation

Ecommerce Europe is the united voice of the European Digital Commerce sector, representing the interests of companies selling goods and services online to consumers in Europe. We welcome the European Data Protection Board's (EDPB) efforts to ensure the clarification of the notion of anonymous data and provide a practical framework for determining whether data has been successfully anonymised. As the EDPB's executive summary states, "data is anonymous if it does not relate to an identified or identifiable natural person" and whether this is the case "may vary from one entity to another." We agree that anonymity should be assessed from each relevant entity's perspective. Kindly find below our observations on targeted aspects of the guidelines.

Introduction - Anonymised data and e-commerce

On SMEs: Anonymised data is part of how small online retailers runs their business. A merchant with a few hundred orders a month uses a reporting app to see which products sell in which regions. Merchants compare their conversion rates with anonymised benchmarks for similar stores. Retailers and payment providers pool anonymised fraud signals to block card testing before it reaches a small shop. None of this needs to identify anyone, and all of it depends on service providers, because an SME cannot build it in-house.

The Guidelines get the fundamentals right. The relative approach to identifiability, based on SRB and Breyer judgment (C-582/14), is the correct reading of the GDPR. The problem is practical. The contextual approach assumes legal and data science resources that most retailers do not have. Several passages also set a bar that, in practice, makes anonymous data hard to reach at all.

Legitimate interest (para 39): The guidelines recognise that anonymisation, especially combined with deleting the original data, is a positive factor when a controller relies on legitimate interest. This rewards good practice and should be welcomed explicitly. We would also ask the EDPB to confirm that anonymisation without deletion of the original data can also count as a positive factor, consistent with para 10, which accepts that anonymisation does not always lead to deletion.

Legacy datasets (fn 3): Datasets already assessed under the 2014 Opinion 05/2014 are not expected to be reassessed. This protects the investment companies have already made. We would appreciate if EDPB states this in the main text as a clear rule, while keeping periodic reassessment as good practice.

Simplified approach (paras 48, 95): The optional simplified approach, and the presumption of anonymity when data passes all three criteria, is the most SME-friendly part of the text. It should be protected in the final version.

Observations on different topics

- **Reassessment over time**

No clear methodology for periodic reassessment (footnote 3)

The EDPB recommends periodic reassessment of anonymisation because technological developments may increase re-identification risks over time, but the Guidelines do not explain how often reassessment should occur, what events should trigger reassessment, whether reassessment applies to all anonymised datasets, whether old datasets must be remediated.

The Guidelines should clarify expectations regarding periodic reassessment of anonymised datasets, including recommended review triggers, frequency, materiality thresholds, and whether reassessment expectations differ depending on the sensitivity of the data and the type of anonymisation used.

Security breach implications require clarification (para 42)

The EDPB suggests that a security incident may require reassessment of anonymity status. It is unclear whether every breach trigger reassessment, only breaches involving auxiliary datasets matter, historical assessments become invalid.

The Guidelines should explain when a security incident requires anonymisation to be reassessed and how this depends on the type and extent of the information compromised.

Drift over time (paras 35, 88, 92-93)

Data that is anonymous today can become identifiable as new data and techniques appear, and the guidelines ask for safety margins and reassessment. A retailer that receives an anonymised benchmark cannot monitor this. Para 29(e) also asks controllers to consider “reasonably foreseeable technological developments”, but a retailer cannot forecast where reidentification research is heading.

Suggestion: “Read para 29 (e) as covering re-identification techniques that are known and available at the time of the assessment, with new techniques picked up at the next reassessment.”

For the proposal below to work in practice, the entity that anonymised the data needs to know who received it, so that it can tell them if the position changes. The proposal is also consistent with footnote 30, which refers to the conditions for controller-to-controller liability set out in Deutsche Wohnen (C-807/21).

Proposed text: *"Reassessment should rest with the entity that anonymised the data. A recipient acting in good faith should not be in breach until it has been notified."*

- **“Means reasonably likely to be used” and residual risk**

"Means reasonably likely to be used" remains difficult to operationalise (para 30)

The EDPB expands the concept and includes investigative journalists; rogue employees; cybercriminals; foreign intelligence agencies; unethical companies. This is problematic because the concept becomes almost limitless as the draft does not establish any practical threshold.

The Guidelines should provide clearer risk-based criteria and examples that define the level of assessment expected for external threat actors such as cybercriminals, investigative journalists and foreign intelligence services, including what evidence is sufficient to demonstrate that re-identification is not reasonably likely.

Suggestion: "Para 9 and footnote 21 already say that an entity with no access to the data, and no link to anyone who has it, does not need to be assessed. Make this the first filter for the actors listed in para 40: a foreign intelligence agency or a cybercriminal matters for a small shop's sales data only if there is a plausible way for the to reach it.

Compelled Access (para 31)

Small online shops typically rely on cloud, payments and analytics providers based outside the EU, or with parent companies there. Para 31 treats an entity being "forced by outside circumstances" to identify individuals as a relevant risk, without a threshold.

Suggestion: "Clarify that compelled disclosure under third-country law is relevant only where there is concrete evidence on a realistic pathway, and only where the compelled entity holds or can obtain the means of re-identification."

Practical impossibility (para 26)

Para 26 gives identification that "would be impossible to do" as one reason why means are not reasonably likely to be used. The Court's case law speaks of identification that is practically impossible (Breyer C-582/14, para 46). For a smaller retailer, the real question is whether anyone could realistically identify a customer from a sales report, not whether it is conceivable in theory.

Suggestion: "use the Court's 'practically impossible' in para 26".

Residual risk (paras 21-22)

Para 22 applies the right test: identification must be "insignificant in reality". Paragraph 21, then treats data as personal if a transfer is likely and it "cannot be ruled out" that the recipient has the means to identify. "Cannot be ruled out" is close to an absolute standard. If paragraph 21 stands as written, almost no shared dataset qualifies as anonymous. The second sentence of paragraph 21 adds that such a transfer makes the data personal for the transferring entity too, even where it could not otherwise identify anyone. The consequence should also follow only where the "insignificant in reality "test is not met.

Suggestion: *"Align para 21 with the 'insignificant in reality' standard in para 22." A merchant that shares minimised sales data with a benchmarking or analytics provider should be judged on what it could reasonably foresee when it shared the data. If the recipient later misuses or leaks it, that should not make the original sharing unlawful after the event".*

- **Contracts, service providers and processors**

Treatment of contractual controls appears ambiguous (para 34)

The EDPB states that contractual prohibitions should not be treated as legal prohibitions and should only complement technical controls. The guidelines say contracts only complement technical measures and must be "reliable, verifiable and enforceable". The text gives no example of what meets the bar. Online retailers typically share data through platforms and app ecosystems, where contracts sit alongside technical controls.

The Guidelines should clarify the role and weighting of contractual, audit and governance controls within the assessment of means reasonably likely to be used, particularly in business-to-business data sharing arrangements.

Suggestion: *Para 33 lets authorities assume that people follow the law unless there is concrete evidence otherwise. Take the same approach to platform terms. Where a ban on re-identification is clear, can be checked and is enforced (for example by suspending access) and sits alongside technical controls, it should count unless there are signs it is being ignored.*

Service providers (para 15, Example 3)

Example 3: an e-commerce retailer sends customer records to an analytics agency, and the retailer's perspective is applied to both parties. That is correct for a raw data transfer. The text should also recognise the opposite case, where the provider only receives a minimised dataset and is technically and contractually barred from re-identifying anyone.

Suggestion: *"Add a counterexample where the processor receives only a minimised dataset, has no access to additional information, and is subject to technical and contractual restrictions on re-identification."*

The basis for this already exists in the guidelines: under paras 9 and 22, data is not personal for an entity that has no realistic way to identify anyone.

A related case matters even more to small retailers. Benchmarking tools and shared fraud-prevention services usually work by combining data from many merchants into one service. A provider running such a service often decides for itself how the pooled data is used, so it is not simply processing on behalf of each merchant. Example 4 suggests that, in that case, anonymity should be assessed from the provider's own position.

Suggestion: "*Add an example of a benchmarking or fraud-prevention provider that combines minimised data from many merchants and assess anonymity from the provider's position*."

Anonymisation by processors (para 38)

The legal basis for anonymising is presumed to be the legal basis of the original processing, as long as anonymisation is part of the same activity and serves the same purpose. SMEs rarely anonymise data themselves; their service provider does it for them.

Suggestion: *"Confirm that the presumption also covers anonymisation carried out by a processor on the controller's instructions. A processor acting on instructions carries out the controller's own processing, so the presumption in para 38 should already apply."*

Anonymising for reporting and analytics (paras 38-39)

Para 38 assumes the legal basis carries ove4r only where anonymisation is part of the same activity and serves the same purpose. In retail, data collected to fulfil orders is often anonymised later to produce sales reports or benchmarks.

Suggestion: "*State that anonymisation to produce reports or analytics will normally pass the Article 6(4) compatibility test, since the output no longer identifies customers*."

Outliers and mixed datasets

Outliers (para 36)

The EDPB states that if anonymisation is effective only for some records, the whole dataset should generally be treated as personal data.

Real-world datasets often contain millions of records, a few outliers, varying re-identification risks. The Guidelines should provide examples of acceptable approaches for handling small numbers of high-risk records or outliers within otherwise effectively anonymised datasets.

Suggestion: "Para 36 treats the whole dataset as personal only where its personal and anonymous parts are not handled separately. Confirm that keeping a record of which records are anonymised, and storing them apart from the rest, is enough to show separate handling."

- ## "Contextual" vs "Simplified Approach"

The "Contextual Approach" vs "Simplified Approach" lacks operational guidance (para 47)

The EDPB introduces two methods: Contextual Approach and Simplified Approach and states that both can be used to assess anonymity. The Guidelines explain the legal theory but do not explain when an organization should choose one approach over the other; whether regulators have a preferred approach; whether one approach creates a presumption of compliance; how assessments should be documented.

The Guidelines should clarify under which circumstances controllers are expected to apply the contextual approach versus the simplified approach. Additional practical examples and minimum documentation requirements would support consistent implementation across organisations and across supervisory authorities.

- ## Aggregate data

Aggregate data treatment requires clarification

(paras 76-81, 86)

The guidelines explain how aggregate data can reveal individuals, especially outliers. They do not say when aggregates are safe. Sales reports and cross-store benchmarks are aggregates, and a small retailer needs a threshold it can apply.

Suggestion: "*Add a worked example using minimum group sizes and suppression of small cells, for instance building on Example 11 (a shop's purchase records).*"

Inference risks in aggregate data (para 74)

The EDPB explicitly recognizes that aggregate data can still relate to individuals and may create inference risks. It refers to membership inference attacks, reconstruction attacks, de-aggregation techniques, and AI-assisted extraction of information from aggregate datasets but does not say when these risks are low enough to disregard.

The Guidelines should provide more operational examples illustrating when aggregate datasets can reasonably be considered anonymous and when additional safeguards are expected to address risks (such as membership inference, reconstruction attacks and AI-assisted de-aggregation)**.**

Leave-one-out test (fn 43; paras 71(b), 77, 82)

Footnote 43 introduces a counterfactual test: would the output change meaningfully if one person's data were removed? This is a useful idea, but there is no guidance on how to apply it.

Suggestion: *"Add a worked example of the leave-one-out test applied to a typical retail or analytics dataset."*

Terminology in privacy notices (para 40)

Para 40 bars “anonymous”, “de-identified” and “de-personalised” where individuals remain identifiable. Online retailers selling across borders often use one privacy notice for several markets. In some of those markets, “de-identified” is a defined legal term for data that remains personal under the GDPR.

Suggestion: *Limit para 40 to statements suggesting that data falls outside the GDPR when it does not and allow “de-identified” where the notice makes clear that the data remains personal data.*”

- ## Digital Omnibus

The Commission's Digital Omnibus proposal amends the definition of personal data in Article 4(1) GDPR. The guidelines do not mention it. Businesses should not have to implement one standard now and a different one later.

Suggestion: *"Commit to reviewing the guidelines once the co-legislators have concluded on Article 4(1). In the meantime, the guidelines should consistently apply the entity-relative approach confirmed in EDPS v SRB (C413/23P).*