A brief analysis of the EDPB Guidelines 02/2026 on anonymisation, with federated learning as a practical example

Joint feedback by FIZ Karlsruhe, Germany and INRIA Sophia Antipolis, France

The guidelines implement the relative approach confirmed in EDPS v SRB: whether information is anonymous depends on the entity concerned and the means it is reasonably likely to use, and the residual risk of identification must be insignificant in reality (paras 7, 22). Three criteria — No Record Isolation, No Linkage and No Inference — serve to evaluate the effectiveness of re-identification (techniques) (para 83), and a violation does not of itself make data personal (para 52). They are described in a simplified approach that disregards differences between entities regarding the means they are reasonably likely to use (para 53), while the contextual approach, which is the legal standard (para 45; fn 36), builds the means test into each criterion (fn 38). Federated learning and analytics make a demanding test case: they are designed so that raw data never leaves the participating sites, and platforms such as Fed-BioMed have been used, for example, to train a cardiac CT model across eight hospitals whose code and weights were then published.

First, the object of the assessment: Federated architecture restricts access to the source data, which remains personal data at each site; it does not determine whether what leaves the site is anonymous. Para 90 makes the general point: technical, organisational and contractual measures can help to restrict access but are not by themselves sufficient to make data anonymous; what must be assessed in more depth is the output— individual updates, per-site and global statistics, the trained model and the analytics query interface. Paras 77 and 82 bring such outputs within scope, but the Glossary definition of "given data" does not name them, and the criteria are framed for records and tables (paras 55, 57, 85). Furthermore, for updates and models, No Record Isolation barely applies. Thus we ask the EDPB to name such outputs in the definition of “given data” and explain how the criteria apply to them and to integrate a more detailed means test.

Second, relativity: Assessing anonymisation per recipient is the guidelines' main strength (paras 7, 12, 22), and a federated deployment shows why it matters, since its outputs reach different recipients. Nevertheless questions remain:

Output	Recipient	Main question
Individual model updates	Central server, unless secure aggregation is used	Could these entities, with means reasonably likely to be used, reconstruct or infer records or attributes of a site’s patients from that site’s update, with enough confidence to attribute them to an individual (para 87)?
Per-site and global statistics	Central server	Could these entities single out an individual from a small site's result, or by comparing repeated queries (paras 76–78)?
Aggregated model	Central server and every participating site	Could these entities, combining the model with a site's own data, draw a specific and meaningful inference about patients at other sites, including their membership (paras 67, 79), also by comparing successive versions?
Published weights	The public	Could any entity with means reasonably likely to be used infer membership or extract training data?

The guidelines do not carry this principle through consistently; e.g. in Annex 1. This specifically concerns the assessment itself and joint controllership, which often arises in multi-site studies. A central server operator that defines the training plan together with the participating hospitals may well be in that position for the local training stage (C-40/17 *Fashion ID*): whether the outputs it receives are anonymous in its hands remains a question of the means. Under para 28, an entity's means include those available to another entity it can draw on, and footnote 7 counts "collusion or cooperation" between entities as a reason to consider them together. Whether joint controllers are to be treated as cooperating in that sense — so that the hospitals' patient data counts among the operator's means — is the decisive question, and the guidelines address neither it nor, more generally, whose perspective governs among joint controllers.

Third, the threshold and the means: The CJEU's standard is that the risk of identification must insignificant in reality (Breyer; EDPS v SRB), and paras 26 (although para 26 drops the Court's "practically" from "practically impossible") and 32 correctly treat legal prohibition and practical impossibility as indications of that result rather than independent exclusions. But the guidelines express the threshold inconsistently: e.g. para 22 speaks of the likelihood of being "distinguished", para 21 treats data as personal where a recipient's means "cannot be ruled out", and para 29(b) asks for "effective and permanent" restrictions on access. In the federated case these formulations decide the outcome. We therefore ask the EDPB to express the threshold consistently, to state what each formulation measures and how they relate to one another, and to set out a framework in which the threshold(s) can be applied — and reviewed — in practice.

Under para 89, appropriate security controls confine the relevant means to authorised users: with secure aggregation the central server receives only the combined update, approval of training code limits what runs on site, and analytics can be disabled. Encrypted transport protects against interception, but encryption limits access only for those without the key and does not make data anonymous for the recipient who decrypts it (para 91). These measures determine whose means count; they do not themselves make the outputs anonymous (para 90). Consortium agreements can complement them, but a contractual prohibition is not a prohibition by law and must be reliable, verifiable and enforceable (para 34). Where re-identification is prohibited by statute, as within a secure processing environment (SPE) under Article 73 EHDS: such a prohibition is a prohibition by law, to which the presumption in para 32 could apply.

Where the model can be passed on, para 21 treats it as personal unless the recipient's means can be ruled out — and then as indirectly personal for those who passed it on. Where weights are published, para 89 treats them as accessible to anybody, so the means of any entity that could realistically obtain them count. Publication does not itself make the model personal data, but it cannot be undone: the assessment must hold, with the safety margins para 88 calls for, before release. We ask the EDPB to provide further explanations, specifically in regard to the federated learning environment. In particular, para 21 should be aligned with para 28; para 29(b) should drop "permanent"; para 34 should say what makes a contractual measure reliable, verifiable and enforceable — for instance, whether SPEs, records of approved code or audit rights suffice; and collusion (fn 7) needs guidance on how it is to be weighed. Where record-level health data fails the criteria, as it ordinarily will, it is this second assessment that decides. Yet it is the less developed part of the guidelines. The

guidelines also leave unclear how the three criteria, the means test and the applicable thresholds relate to one another, and whether they apply in isolation or in combination. We therefore ask the EDPB to clarify the definitions used in the criteria and to treat the means test as a distinct step preceding the application of the criteria, to provide a second-stage analysis for No Inference that weighs the means of each relevant entity, and to relate formal guarantees such as differential privacy to para 22.

Fourth, the criteria: For federated outputs No Inference decides most cases: reconstruction from updates, membership inference on the model and differencing on per-site statistics are all inference attacks (paras 67, 76–79, 82). Two weaknesses meet here. A No Inference violation leads directly to "likely to be personal", without the second-stage analysis properly provided for the other criteria (para 97; Annex 1). And differential privacy earns no clear credit: footnote 43 uses the leave-one-out comparison on which it rests, but its parameters are never related to para 22 and the guidelines name no acceptable level. Para 87 and Example 20 nonetheless supply the right test — whether a result can be attributed to a particular individual with sufficient confidence. Nevertheless, the distinction between the three criteria, the means test and the applicable thresholds is unclear, as is the extent to which they apply in isolation or in combination. Therefore, we request clarification of the definitions and thresholds, a clear framework and standards that can answer the pertinent question of how to assess residual reidentification risk. The threshold should be expressed consistently as the risk of identification. Para 21 should be aligned with para 28. Para 29(b) should replace "permanent". Para 34 should say what makes a contractual measure reliable, verifiable and enforceable — for instance, whether SPEs, records of approved code or audit rights suffice. And collusion (fn 7) needs guidance on how it is to be weighed.

Fifth, the aggregation rules: Under para 36, read with para 87 and footnote 46, one identifiable individual can make a whole output personal: a patient with a rare condition at a small site may taint an update or that site's statistics. And every training round or analytics query is a further output from the same data, yet the guidelines discuss multiple outputs mainly as an attacker's method (para 77) and reassessment as good practice (para 35). Where a federated analytics set-up lets the central server request each site’s statistics separately — say, the number of patients and their mean age at each hospital — and not only the combined figure, a site with only one or two patients matching a query in effect discloses an individual’s values. Unless results below a minimum number of records are withheld, the output may well be personal data for its recipient. Output checking in trusted research environments addresses exactly this, looking at small numbers, rare characteristics, outliers and unusual combinations of attributes.

We agree with the guidelines' overall approach but their operative parts lack in clarity and were mainly written for tables released once. Federated settings, in which models and statistics are released repeatedly to several recipients, expose that gap. We ask the EDPB for a worked example on federated learning and analytics: a model trained across several hospitals, with its individual updates, per-site statistics, aggregated versions and published weights assessed for two or three recipients with different means following the steps from access through inference (reconstruction, membership inference, collusion between parties) to identification, and showing what evidence supports each step, such as the secure aggregation set-up, the differential privacy

parameters and the results of membership-inference testing. We further ask for: an express statement in the Glossary that “given data” may consist of model updates, trained models, aggregate statistics and query interfaces (paras 77, 82); an explanation of how differential privacy and similar formal guarantees, and their parameters, relate to the requirement that the risk be insignificant in reality (para 22; fn 43); the assessment of each recipient’s means set out as a distinct step, with examples as detailed as those given for the criteria; a requirement that outputs derived from the same data — successive rounds, repeated queries, separate releases — be assessed together; and guidance on whose perspective governs where controllers act jointly (Article 26), as in multi-site studies.