Comments on the EDPB Guidelines 02/2026 on Anonymisation from a medical-technology perspective

Contribution to the public consultation of the European Data Protection Board · October 2026

I. Introduction and general remarks

Siemens Healthineers thanks the European Data Protection Board for the opportunity to comment on the Guidelines 02/2026 on anonymization, adopted for consultation on 7 July 2026. With the Guidelines the Board updates the aging WP29 Opinion 05/2014. The trigger is familiar: evolving case law, new data spaces, and above all the advances in AI and re-identification. For a medical-technology company whose value creation increasingly rests on clinical imaging, genomic and real-world data, this is not a marginal topic. Whether a dataset counts as anonymous determines whether the GDPR applies at all, and thus the usability of entire data holdings for product development, AI training, post-market surveillance and research.

The good news first. The Board clearly embraces the relative, context-dependent notion of anonymity. Anonymity is assessed from the perspective of the entity concerned, and data may be personal for one entity and anonymous for another (paras 6–7 and footnote 4, referring to C-413/23 P EDPS v SRB, where the Court held that pseudonymized data are not to be regarded as personal data in all cases and for every person, para 86 of the judgment). This reflects the reality of distributed data ecosystems and deserves support.

The downside is serious. The technical test (No Record Isolation, No Linkage, No Inference) bears down hardest on high-resolution, high-dimensional data at the level of individual records (paras 85–86). Yet that is exactly what medical cross-sectional images, whole-slide pathology and omics data are; they are by nature almost always unique and therefore regularly fail the first criterion already. Combined with the warning about AI-driven re-identification, including agentic AI (para 92), the time-axis logic (para 35) and the downgrading of purely contractual safeguards (paras 34, 90), this threatens a de facto prohibition on anonymizing the most data-intensive product classes.

Decisive is a point the Guidelines themselves lay down: uniqueness is not the same as attributability. A high-resolution individual record may fail the first criterion and still be anonymous if singling out cannot be achieved for lack of matchable additional information, or if the unique record is not intelligible, provided the other two criteria are met (para 99, Example 24). This dividing line determines whether the framework is fit for medical technology.

Nine comments carry these observations. Two of them (Comment 2 and Comment 3) go to the core of data-driven medical technology and are marked as key concerns; the others aim at legal certainty without lowering the level of protection. Medical technology is an already highly regulated

ecosystem (MDR, IVDR, AI Act, EHDS, professional and social-law confidentiality) whose sector-specific safeguards the Board should take into account.

II. Background and classification

When may a medical-technology company treat a clinical dataset as anonymous, so that the GDPR no longer applies? The Guidelines answer this question in two steps, a legal analysis (Chapter 2) and a technical test (Chapter 3).

1. The relative notion of anonymity

The core test is two-stage: does the information relate to a natural person (para 16), and is that person identifiable by means reasonably likely to be used (para 18)? Both questions may be answered differently depending on the entity (paras 6–7, CJEU EDPS/SRB). What matters is not absolute but reasonably expected identifiability; the likelihood need not be zero, it must be insignificant in reality (para 22).

2. Means reasonably likely to be used

Whether a means is reasonably likely to be used depends on objective factors (para 29): properties of the data, context and access restrictions, availability of additional information, cost and time, and available and foreseeable technology. The Board makes clear that data under appropriate security controls constitute an access means for authorized users only, not for unauthorized actors (para 89).

3. The technical test

Three criteria carry the test: No Record Isolation (para 55), No Linkage (para 60), No Inference (para 67). If a dataset fails No Record Isolation, one must check whether singling out is possible by matching against attributes (para 98). If singling out cannot be done and the other two criteria are met, the data remain anonymous although unique (para 99). Example 23 affirms identifiability where demographic attributes are matchable via public registers; for the pixel or radiomic features of a medical image, a comparable generally accessible reference source may be lacking, which is precisely what needs to be assessed in the specific case. Example 24 confirms: a dataset may be anonymous despite failing No Record Isolation.

III. Detailed comments

The following nine comments follow the scheme A (summary), B (passage concerned), C (assessment), D (proposed wording).

Comment 1 · The relative notion of anonymity: guard against a relapse into a quasi-absolute standard

A · SUMMARY The relativization by perspective is appropriate. The broadly framed list of relevant entities, extending to cybercriminals and intelligence services, threatens to flatten the relativization in practice if it is applied as a checklist.

B · PASSAGE Paras 6–7, 9 and 11 (perspectives), para 30 with footnote 25 (catalogue of relevant entities), paras 32–33 (rebutting the assumption of lawful conduct), para 89 (access reservation) and Example 22.

C · ASSESSMENT If the circle of relevant entities is routinely drawn out to malicious actors with exceptional capabilities, the relative standard derived from SRB is emptied. The reference to intelligence services is the clearest example: such actors may possess far-reaching capabilities, but whether they have any realistic access to a given dataset is a question of the specific case. The Guidelines themselves point in this direction: footnote 25 states that not all of the listed perspectives need to be considered systematically in each case; under para 89, access to appropriately secured data is a means reasonably likely to be used only for authorized users; and in Example 22 persons who are extremely unlikely to receive the data are not treated as relevant entities. At the same time, Example 22 counts persons who might hack into the system and public authorities as relevant entities in view of a possible leak. The catalogue in para 30 therefore needs a clear rule on when an entity is to be included, rather than a list that is read as a default.

D · PROPOSAL The Board should move the statement in footnote 25 into the body of para 30 and clarify that the list of possible entities creates no presumption of their relevance. An entity is to be included where it has, or is likely to obtain, access under paras 9 and 89 or is part of a likely access chain. For actors who are not effectively bound by Union or Member State law, such as foreign intelligence agencies, inclusion should rest on concrete indications of a risk of access, by analogy with the evidence-based approach of para 33, rather than on the merely abstract possibility of exceptional capabilities. For public authorities with legal powers of access, the assessment under para 29 should refer to the concrete legal framework governing access to the dataset in question.

Comment 2 · High-resolution, high-dimensional health data: a threatened de facto prohibition on anonymization KEY CONCERN

A · SUMMARY Medical cross-sectional images, digital pathology and omics data are by nature unique at record level and regularly fail No Record Isolation. Without clarification the test threatens to exclude these data classes permanently from anonymity. At the same time the Guidelines lay down the way out: uniqueness is not attributability.

B · PASSAGE Paras 55–57 (No Record Isolation), paras 85–86 (re-identification most effective against high-dimensional, high-resolution records; effectiveness depends on the availability of

additional information), paras 98–99 and Examples 23/24 (singling out; uniqueness despite absence of attributability), reinforced by para 92 (AI-driven re-identification).

C · ASSESSMENT Example 10 already shows uniqueness at four attributes; a volumetric dataset or a genome contains many times as much. Applied mechanically, virtually every imaging or genomics cohort would have to be treated as personal, with full GDPR and Article 9 binding even for pure model development. This goes to the core of medtech AI value creation and sits uneasily with the secondary use of health data that the European Health Data Space is designed to enable. The Guidelines, however, do not sufficiently acknowledge that data type and personal reference diverge: high dimensionality and resolution raise the intrinsic re-identification potential (paras 85–86) but do not by themselves establish personal reference. Comprehensive genomic data may have a particularly high singling-out capacity; a narrowly bounded organ image may be assessed differently in the recipient context if it can be shown empirically that neither pixel and radiomic features nor available reference data allow singling out, linkage or a specific and meaningful inference about an identifiable patient. Precisely this distinction is provided by para 99: a unique record remains anonymous, provided the other two criteria are met, if its attributes cannot be matched to other available data or the unique record is not intelligible (Example 24). For head scans, the singling-out test and No Linkage deserve particular attention, since facial features rendered from the scan could be matched against photographs (paras 60 and 98); this is exactly what domain-specific techniques such as defacing address. In practice, two ways out of a No Record Isolation violation follow: first, showing that singling out is not possible because additional information is missing or the data fields are sufficiently reduced (Example 23 a contrario); second, aggregation and, depending on the case, generalization, against which re-identification techniques are less likely to be effective (paras 58 and 86).

D · PROPOSAL The Board should add a dedicated health-data example separating the uniqueness of a high-resolution record (e.g. a defaced MRI) from actual attributability, and recognize that domain-specific state-of-the-art techniques (defacing, generalized metadata, differential privacy, synthetic data, aggregation) can render high-resolution health data anonymous for a clearly delimited recipient environment, provided singling out under para 99 is genuinely excluded and the other two criteria are met. Federated learning, which avoids sharing the data in the first place, should be acknowledged as a complementary approach. The Board should clarify that a No Record Isolation violation alone does not establish personal reference, but that the singling-out test under paras 98–99 is determinative.

Comment 3 · An Article 9(2) basis for the anonymization operation itself: blocking the risk-reducing step KEY CONCERN

A · SUMMARY Under para 38, anonymization, as processing in its own right, requires a legal basis and, for special categories, an Article 9(2) exemption. For the downstream anonymization of existing clinical data, controllers may find it difficult to identify a fitting exemption, so that the very privacy-friendly step is legally obstructed.

B · PASSAGE Para 38 with footnote 33 (Article 6 and Article 9(2); presumption of alignment only for the same activity and purpose; reference to Guidelines 1/2026), flanked by para 39 (anonymization, at least combined with deletion of the original data, as a positive factor in the balancing of interests).

C · ASSESSMENT The alignment presumption in para 38 is welcome but narrow: it applies only where anonymization is part of the same processing activity and serves the same purposes. Where a controller anonymizes a historical clinical dataset at a later stage, for example to make it available for AI development, activity and purpose will often not be the same. For health and genomic data, controllers may then have to rely on research provisions (Article 9(2)(j) together with national rules such as Section 27 of the German Federal Data Protection Act) or other narrow exemptions. The paradox arises that the step which reduces the risk to data subjects may fail at Article 9. This is hard to reconcile with Article 25. Anonymization is, in its effect, the functional equivalent of deletion: it severs the link to the individual and serves data protection by design (Article 25), and should therefore not be subject to stricter requirements than the continued storage of identifiable data. The same tension appears in para 39, which ties the positive effect of anonymization in the balancing of interests to the deletion of the original data, although in healthcare the source data must regularly be kept for statutory retention periods; para 10 itself recognizes that anonymization within the meaning of the Guidelines does not necessarily lead to the deletion of the original data. The Guidelines already point to a route: footnote 33 refers to Guidelines 1/2026, and anonymization can be understood as compatible further processing (Articles 5(1)(b) and 6(4)). The issue is nevertheless the clearest case for clarification beyond mere interpretation of the Guidelines and is part of the ongoing legislative debate on the Digital Omnibus. As long as a legislative clarification is pending, the Guidelines should not settle the question in a restrictive sense.

D · PROPOSAL The Board should recognize that effective anonymization which serves to reduce risk is a privacy-enhancing safeguard applied to lawfully processed data and does not require a legal basis or an Article 9(2) exemption of its own; the subsequent use of the anonymous output falls outside the GDPR, and the fact that the output is subsequently used, for example for research or AI development, does not turn the anonymization itself into a use of the personal data for another purpose. At the very least, the Board should recognize that anonymization of special categories, as a data-protection-by-design measure, can rely on the same Article 9(2) exemption as the preceding processing where the anonymization constitutes compatible further processing, align this expressly with Guidelines 1/2026, and in case of doubt choose a risk-reducing interpretation. It should also clarify that the deletion of the source data is not a condition for recognizing the risk-reducing effect of anonymization, in particular where statutory retention obligations apply. Where interpretation reaches statutory limits, Siemens Healthineers encourages the Board to support a legislative clarification in the context of the Digital Omnibus and the EHDS.

Comment 4 · Contractual and organizational measures: given too little weight

A · SUMMARY In secure processing environments enforceable contractual and organizational controls are the backbone of limiting the means reasonably likely to be used. Their blanket downgrading undermines realistic anonymization in regulated ecosystems.

B · PASSAGE Para 34 (a contractual prohibition is not a legal prohibition), para 89 (access under security controls only for authorized users) and para 90 (technical, organizational and contractual measures are “not by themselves sufficient”).

C · ASSESSMENT That contracts alone do not suffice is true and not disputed. But the blanket subordination misjudges the reality of closed processing environments, such as the secure processing environments provided for in Article 73 of the European Health Data Space Regulation, in which access is restricted to authorized persons: binding, audited and enforceable access regimes reduce the likelihood of re-identification in a real and verifiable way, and the Board itself acknowledges in para 89 that security controls limit the circle of those with access. What must be avoided is a blanket suspicion that contractual arrangements are made only for show, not observed, or set aside when convenient. The remark that contractual measures “may even be disregarded entirely by the parties” (para 34) should not turn into such a presumption; the Guidelines point to no evidence that this occurs in regulated environments, and contractual controls are regularly combined with technical ones. The Board itself names the right criteria in para 34: the measures should be reliable, verifiable and enforceable. The Court of Justice has likewise accepted that technical and organizational measures, in that case pseudonymization, may have the effect that, for a recipient, the data are not personal in nature, provided the recipient is not in a position to lift them and the measures in fact prevent attribution, including by cross-checking with other factors (C-413/23 P EDPS v SRB, para 77).

D · PROPOSAL The Board should clarify in para 90 that, in a secure processing environment, binding, verifiable and enforceable contractual and organizational restrictions, combined with technical measures, can be decisive under the contextual approach for finding that certain means are not reasonably likely to be used, and that such arrangements are to be given credence absent concrete indications to the contrary. Consistency with the secure processing environments of the EHDS should be ensured. The same applies to binding and enforced intra-group rules, which provide a high degree of assurance that commitments will be observed; such safeguards should be assessed for their actual reliability against the criteria of para 34.

Comment 5 · Time axis and technological progress: open-ended liability without a reliable assessment point

A · SUMMARY The requirement to take account of foreseeable developments including agentic AI, together with re-personalization over time, creates open-ended liability for long-lifecycle products. What is needed is a periodic, risk-based review without retroactive illegality.

B · PASSAGE Para 29(e) (foreseeable developments), para 35 with footnote 30 (rising likelihood over time, relapse into personal reference, continuing responsibility), para 88 (safety margins) and paras 92–93 (AI and agentic AI reduce effort and cost).

C · ASSESSMENT Medical devices and their AI components often remain in use for many years, and post-market surveillance under the MDR and IVDR accompanies them throughout their lifetime. A duty to price in speculative future capabilities today is neither feasible nor proportionate and creates a moving target; an overly broad standard would generate documentation requirements without real protective value. What is needed is a workable standard of what is reasonably foreseeable, combined with periodic, risk-based reviews. The key clarification: para 35 provides that data "should again be considered personal" once the likelihood of identification is no longer insignificant. This should be confirmed to operate only for the future, so that a careful assessment made and documented according to the state of the art at the time of anonymization does not render the earlier processing unlawful. The reference in footnote 30 to the case law on controller liability (C-807/21 Deutsche Wohnen) sharpens the sanction perspective and makes the clarification pressing.

D · PROPOSAL The Board should clarify that (a) "foreseeable" within para 29(e) captures only concretely foreseeable, not merely theoretically conceivable developments, (b) consistent with para 35, a relapse into personal reference operates only for the future, so that an assessment made carefully and documented at the time of anonymization does not render the earlier processing unlawful, (c) reassessment, which para 35 treats as good practice, is framed as a periodic, risk-based review, supplemented by reassessment upon material changes, such as new additional information, a security incident (para 42) or a material technological development, and (d) after a transfer, each recipient reassesses from its own perspective, and a recipient which relied on a careful, documented assessment does not retroactively act unlawfully when circumstances later change.

Comment 6 · No Inference for AI models and synthetic data: the boundary is too blurred and the consequences underestimated

A · SUMMARY Trained models and synthetic datasets can become personal via No Inference. The boundary between permissible generalization and impermissible, record-bound inference is not sufficiently operationalized for models; the consequences of a precautionary classification are not addressed.

B · PASSAGE Paras 67–71 (specific and meaningful inference; delimitation from population knowledge; ways of demonstrating reliance on the records in para 71(a) to (c) with footnotes 43 and 44), para 82 (inference from aggregate data, AI models and synthetic data) and para 77 with footnote 45 (cross-reference to Opinion 28/2024).

C · ASSESSMENT The pairing specific and meaningful is convincing in principle, and the Guidelines cleanly delimit population knowledge (para 70). Para 71 also offers ways to show whether an inference relies on the records of an individual: the comparison of outcomes with and

without those records (para 71(b)) and a rebuttable presumption where the inference could have been drawn from general knowledge (para 71(c)), with footnote 44 requiring a significantly different "posterior belief". For trained clinical models, however, these tests still need to be translated into practice: a comparison under para 71(b) would in principle require training without the records concerned, which can be very burdensome for large models, and the threshold in footnote 44 is not yet linked to established technical evidence. Without such operational guidance, every model trained on clinical data risks being treated as personal as a precaution. This classification has significant, unaddressed consequences: it would not only impede validation with synthetic data and the sharing of models, but also trigger data-subject rights, in particular access and erasure requests against the model, which are barely manageable technically and organizationally and would entail disproportionate effort.

D · PROPOSAL The Board should dovetail the Guidelines with Opinion 28/2024 and operationalize para 71(b) and (c), together with footnote 44, for trained models by recognizing practicable evidence that distinguishes a memorizing, training-record-bound model from a generalizing one, e.g. resistance to membership inference against recognized thresholds as a way of demonstrating that no significantly different posterior belief arises; clarify that the merely abstract risk of model extraction does not breach No Inference; proportionately contain the consequences of a classification as personal (access and erasure requests); and recognize that technical restrictions on access to a deployed model, such as controlled interfaces, monitoring and rate limits, form part of the contextual assessment rather than the model being assessed in isolation.

Comment 7 · Transition from WP29 05/2014 and legacy protection: welcomed but to be firmed up

A · SUMMARY That holdings assessed under the old standard need not be reassessed is welcome. For the large legacy and registry holdings of medical technology, express transition and legacy-protection clarity and an adequate lead time are needed.

B · PASSAGE Footnote 3 (a controller that has assessed a dataset as anonymous under Opinion 05/2014 is not expected to conduct a new assessment; periodic reassessment only as "good practice"), together with para 35.

C · ASSESSMENT Legacy protection is of practical importance but remains softly worded. Supervisory authorities could make the new standard a benchmark early in practice. An express transition framework, an adequate period, and clarification of the interplay with Guidelines 1/2026 and the EHDS create planning certainty.

D · PROPOSAL The Board should raise the legacy protection from the footnote into the body of the Guidelines, provide an adequate lead time before the new framework is used as a supervisory benchmark, and expressly clarify the relationship with Guidelines 1/2026 and EHDS secondary use.

Comment 8 · The “insignificant in reality” threshold and the documented risk decision

A · SUMMARY The Guidelines require the likelihood of identification to be “insignificant in reality” but give no reliable indications of when this threshold is met. In practice the controller must therefore make a risk-based assessment under the accountability principle and document it.

B · PASSAGE Para 22 (the likelihood need not be zero but must be insignificant in reality) together with the objective factors of para 29, the documentation requirement of para 41 and the accountability principle (Art. 5(2) GDPR).

C · ASSESSMENT The standard is appropriate because it does not require absolute certainty. Without guidance criteria, however, it remains open when the threshold is met. In practice it follows that the classification of data as anonymous rests on a risk-based assessment according to the state of the art which, as para 41 requires, is documented so as to demonstrate both the compliance of the anonymization process and the effectiveness of the anonymization itself. The Board should expressly recognize this documented, risk-based approach and, so far as possible, offer guidance criteria.

D · PROPOSAL The Board should (a) provide guidance criteria or standard examples for when the likelihood of identification is “insignificant in reality”, (b) recognize that an anonymity assessment made carefully according to the state of the art and documented in line with para 41 meets the legal requirements and is not devalued by merely abstract residual risks, (c) clarify that a supervisory authority which doubts the anonymity of a dataset should identify the concrete, realistic means or actors it considers reasonably likely to lead to identification, and that a complaint calls for an examination of concrete indications of identifiability without prejudice to the controller’s accountability under Article 5(2) GDPR, and (d) confirm that the documentation under para 41 is proportionate to the risk and may be integrated into existing records, such as records of processing activities, data protection impact assessments and, for manufacturers, the quality management and risk management documentation maintained under the MDR and IVDR.

Comment 9 · Processor perspective and change of role: when the manufacturer’s own perspective applies

A · SUMMARY Medical-technology manufacturers may act as processors for healthcare providers (for example in remote service or cloud-based applications) and, in separate activities, as independent recipients of data. Para 15 rightly aligns the processor’s perspective with that of the controller. What is missing is guidance on the change of role and on how the transfer logic of para 21 applies in such settings.

B · PASSAGE Para 15 with Examples 3 and 4 (perspective of the processor and of an independent recipient), para 21 (data remain personal where an onward transfer is reasonably

likely and identification by the recipient cannot be ruled out) and para 9 (entities without real access).

C · ASSESSMENT The approach in para 15 reflects the allocation of responsibilities under the GDPR and is not disputed. For a manufacturer it is decisive, however, that data which a healthcare provider anonymizes itself, or has anonymized by a processor on its instructions, and then discloses to the manufacturer for the manufacturer's own purposes, are assessed from the manufacturer's own perspective where, as in Example 4, the manufacturer acts independently, determines the means and purposes of its processing, does not act on the provider's instructions and the data are not returned to the provider. Whether such a disclosure is permissible is a question for the controller's decision and its legal basis; it is not a reason to bind the manufacturer permanently to the controller's perspective. Para 21 should be read together with paras 9 and 22. As to condition (a), binding and enforced restrictions on onward transfer bear on whether an entity can reasonably be expected to transfer the data. As to condition (b), the words "cannot be ruled out" reflect the Court's reasoning on subsequent transfers (C-413/23 P EDPS v SRB, para 85), but they refer to identification through means reasonably likely to be used; they should therefore be read consistently with para 22, under which the likelihood need not be zero but must be insignificant in reality, and with the Court's finding that pseudonymization may effectively prevent identification by persons other than the controller (para 86 of the judgment). Otherwise data would remain personal along an entire chain merely because identification by a later recipient cannot be excluded in theory.

D · PROPOSAL The Board should (a) add a healthcare example showing that, where a controller, or a processor acting on its instructions, anonymizes data and discloses the result to a manufacturer acting independently, without the data being returned to the provider, anonymity is assessed from the manufacturer's own perspective (Example 4), (b) clarify in para 21 that binding and enforced restrictions on onward transfer are taken into account under condition (a), and that condition (b) applies the standard of para 22 rather than requiring identification to be excluded in absolute terms, and (c) add examples of controller-to-controller transfers showing which party is expected to perform or verify the anonymity assessment and how a change in processing roles affects the applicable perspective.

IV. Concluding remarks

The Board presents a considered, doctrinally coherent framework. Adoption of the relative notion of anonymity deserves support, as does the separation of the simplified and contextual approaches and the candid naming of the AI risk. The friction lies not in the system but in its application to the most data-intensive areas of medical technology. Where high-resolution individual records are the norm, the boundary shifts far towards personal reference unless the singling-out test under paras 98–99 consistently distinguishes uniqueness from attributability.

These comments are not meant to resist the framework but to contribute to making it workable for medical technology. Of the two key concerns, Comment 2 can be resolved with a dedicated health-data example, whereas Comment 3 calls for a statement of principle that risk-reducing

anonymization requires no legal basis of its own (at the very least, its recognition as compatible further processing) and, where interpretation reaches statutory limits, support for a legislative clarification; the remaining comments, including the role-related clarification in Comment 9, create legal certainty without lowering the level of protection. Coherence with Guidelines 1/2026, Opinion 28/2024 and secondary use under the EHDS is of particular importance, as is consistency with the Data Act and the AI Act, so that the same data are not non-personal under those acts yet personal under the GDPR as read by the Guidelines.

Beyond this lies the wider significance: access to large, high-quality health data is a precondition for progress in the treatment of many diseases and for technical progress under tight budgets; at the same time it is a question of European competitiveness, since AI training and research and development may otherwise shift to other regions. Siemens Healthineers remains at the Board's disposal for any questions and further technical exchange.

V. Sources and citation note

EDPB, Guidelines 02/2026 on Anonymisation, version 1.0, adopted 7 July 2026 (version for public consultation). Paragraph, footnote and example numbers refer to the English language version.

Case law: Court of Justice, judgment of 4 September 2025, C-413/23 P EDPS v SRB, ECLI:EU:C:2025:645, paras 77, 85 and 86. Legislation: Regulation (EU) 2025/327 on the European Health Data Space, Article 73.

Reference documents (as cited in the Guidelines): EDPB Guidelines 1/2026 on processing of personal data for scientific research purposes; EDPB Opinion 28/2024 on AI models; WP29 Opinion 05/2014 (updated by the Guidelines).