30th of October 2026

Novo's feedback to EDPB's public consultation on Guidelines 2/2026 on anonymisation

Novo welcomes the opportunity to comment on the European Data Protection Board's draft Guidelines 2/2026 on anonymisation. Novo is a global pharmaceutical company with headquarters in Denmark.

As a research organisation, Novo relies on health data to advance science, innovate and develop new treatments that benefits patients. This includes collaborating with research partners and sharing and receiving health data. To enable this a clear and workable interpretation of anonymisation is essential. Consequently, the Guidelines should provide a dependable and foreseeable legal standard based on concrete means reasonably likely to be used, rather than every identification scenario that can theoretically be conceived.

Novo welcomes several of the principles presented in the Guidelines. These includes the recognition that the same data can be anonymous for one entity and personal data for another, that anonymisation does not require zero risk of identification, and that anonymisation does not require deletion of the original data.

However, Novo's principal concern is that the technical framework and examples do not consistently reflect the Guidelines' contextual legal approach. Para 28 correctly excludes purely hypothetical combinations of means. Other passages, however, ask what "can" occur, treat unknown capabilities of identification as relevant or assume that safeguards will be circumvented. Requiring controllers to consider hypothetical recipients, unknown datasets or speculative future attacks risks reintroducing the absolute standard that the Guidelines expressly reject. This would limit the ability to conduct scientific research in Europe.

An effective assessment of anonymisation should consider data transformation together with the technical, organisational and legal safeguards that govern access and use. These include statutory and professional confidentiality obligations, enforceable prohibitions on re-identification, audit rights, access restrictions and onward transfer controls. Their combined practical effect should determine which identification means are reasonably likely to be used.

Therefore, we recommend that the final Guidelines preserve a workable, evidence-based route to anonymisation by:

a) recognising established scientific and industry practice for the identification risk assessment, such as ISO/IEC 27559 and the risk-based approach applied under EMA Policy 0070,
b) giving the combined practical effect of technical, organisational and legal safeguards proper weight,
c) distinguishing the ability to single out a random record in a dataset from an identifiable person, and
d) enabling reusable assessments for similar data types with proportionate review over time.

Furthermore, the final Guidelines should allow organisations to assess anonymisation at the level of a processing activity, rather than case by case for every single dataset. For example, one general assessment could cover clinical trials within one therapeutic area, while another could cover the use of real-world data within a different therapeutic area. Each assessment would reuse validated anonymisation methods and documented reasoning while the underlying assumptions remain the same, with additional checks only where a dataset differs materially. Requiring a separate assessment for every single dataset would create a significant additional administrative burden for the life sciences industry and seriously impact the ability to reuse health data for research and innovation.

We also encourage the EDPB to add real-life examples from pharmaceutical research, including clinical trial data, registries, research collaborations and the European Health Data Space. These examples should distinguish unrestricted publication from controlled access and show how safeguards affect the assessment.

More specifically, Novo has identified eight areas where we recommend additional clarification to support responsible research and innovation in Europe. These are covered below, followed by specific proposals to the text and a concluding recommendation.

Consolidated main concerns

a) Legal basis and types of re-identification

- Para 38 presumes that the legal basis and sensitive data exemption for anonymisation coincide with those of the preceding processing, where anonymisation is part of the same activity and purpose. The Guidelines should explain how this relates to Article 89 GDPR safeguards for scientific research, particularly where anonymisation is itself a required safeguard.

- Additionally, the presumption in para 38 does not cover anonymisation undertaken to prepare existing personal data for a different purpose. We encourage the EDPB to explain how the legal basis requirement applies in that situation (see also our comment on para 38 below). Without a simple and clear route to reuse personal data for other purposes by way of anonymisation, data that already exists may need to be re-collected solely because the legal basis for further use is unclear. If a simple and clear ability to reuse health data for secondary purposes does not exist, it will inevitably lead to more medical interventions to re-collect data that already exists but where the legal basis for further use is unclear.

- Para 87 and Example 20 rightly require a re-identification technique to produce results with sufficient accuracy and confidence to distinguish an individual and treat them differently. The Guidelines should apply this standard consistently across the technical assessment. In particular, the Guidelines should explain how confidence, the ability to attach a result to a person and its consequential use should influence safeguards. An uncertain match that cannot reliably be attached to a person should not be treated as identification.

- A later accidental identification should not, by itself, show that an earlier anonymisation assessment was wrong. Para 27 states that information is not anonymous once an individual has been identified, and para 31 recognises that identification can happen by accident or negligence. The Guidelines should clarify that para 27 concerns the status of the data from that point onwards. Recital 26 asks whether identification is reasonably likely, so the original assessment should be judged against the means reasonably foreseeable at the time. Under para 22, the likelihood of identification must be insignificant, not zero, so an unlikely event can still occur. When it does, the controller should respond and reassess, in line with para 35. If every accidental identification proved that data was identifiable, the zero-risk standard that the Guidelines reject would effectively return.

b) Contextual approach, entity perspective and processing chains

- The term “entity” and the selection of relevant perspectives need greater precision. The Guidelines should explicitly address whether and when functions or departments within the same legal entity may receive data that is anonymous from their operational perspective, even though another part of the organisation retains identifying information. For example, where a psychologist is bound by professional confidentiality, their patient records are not accessible to others in the same organisation. Anonymised case data used for research across that organisation therefore cannot easily be re-identified merely because the psychologist holds the original record.

- We agree with para 15 and Example 3 that a processor cannot treat data it processes on behalf of a controller as anonymous where that data is personal data for the controller. The controller determines the purposes and means of the processing and is therefore the operative entity for the assessment. We recommend that the Guidelines ground this conclusion more explicitly in the processor's legal position under Articles 28 and 29 GDPR. A processor acts only on the controller's documented instructions and under its authority, so it cannot disregard identification means attributable to that controller. Where a recipient instead determines its own purposes, it acts as a controller, and its own perspective should be assessed under the contextual approach.

- Para 21 may cause otherwise anonymous data to be treated as personal data through extended chains of possible transfer and third-party access. We request a limiting principle tied to realistic access such as credible onward-access routes (taking account of contractual restrictions), evidence of likely access, connection between datasets that would enable re-identification, and means reasonably likely to be used. We also request that "cannot be ruled out" language is reconciled with the para. 22 test of "insignificant in reality".

- We welcome the clarification that the simplified approach is voluntary and is not the only means to demonstrate anonymity. Controllers may rely directly on a properly documented contextual assessment, and failure to satisfy the simplified approach does not establish that the information is personal data. However, describing that approach as providing "greater confidence" could encourage organisations or authorities to treat it as the preferred benchmark, thereby shifting the practical burden of proof and creating a presumption that information remains personal unless the simplified criteria are met. In research settings, this could lead to unnecessary transformation or withholding of data.

c) Identifiability must be more than uniqueness

- Example 5, involving a person wearing a suit and walking a dog, risks being read as equating singling out within a specific context with identification of an individual. We request that the EDPB clarify that an additional route to identity is ordinarily needed, such as a visible face, trackable movement, or matchable secondary data from other sources. Record isolation should remain a screening criterion, not a conclusion. The Guidelines should foreground the later qualification in paras 98-100 that isolated records must be matchable to a real individual.

- Personal attributes that change over time, such as glucose levels, blood measurements or cognitive signals, should not be treated as stable quasi-identifiers without assessing

temporal persistence and realistic repeat measurement. We recommend that the Guidelines include examples clarifying the difference between stable quasi-identifiers and values that change over time.

d) A bounded and operational test for "means reasonably likely to be used"

- The draft does not provide enough practical guidance on what evidence demonstrates that a re-identification route is realistic rather than merely conceivable. The assessment should principally consider the lawful, concrete and foreseeable means available to the relevant entity, not means theoretically held by unidentified third parties.

- The Guidelines should expressly consider: the number and type of capable actors; access to the dataset and secondary sources or information; technical capability; time, cost and labour; likely benefit; enforcement risk; whether the data changes over time; and whether the same information could be obtained more simply from another source.

- The discussion of current re-identification techniques may create a continuously expanding scope without a safe harbour or clear review triggers. Safety margins are useful but should not turn hypothetical future capability into present identifiability that would prevent data from being anonymised and used for research.

e) Contracts, professional confidentiality and security controls

- Para 34 gives insufficient weight to contractual restrictions. We do not argue that a contract alone automatically produces anonymity, but ask the EDPB to recognise reliable, verifiable and enforceable re-identification prohibitions, audit rights and onward-transfer controls as meaningful mitigating factors.

- The guidance should also address legal and professional confidentiality regimes that prevent re-identification or access, including clinical confidentiality, rather than focusing mainly on laws that enable access.

- Rogue-employee and malicious-actor scenarios should be included only where realistic access and concrete evidence justify their inclusion, without assuming that enforceable duties and controls in place will fail.

- The Guidelines should distinguish assessment of identifiability from planning for accidental or unlawful access. The inability to guarantee that a security incident will never occur should not defeat anonymity. A successful attack may expose identifying information and create a personal data breach requiring containment, assessment and follow-

up actions. It does not necessarily establish that the attacked dataset was personal data before that event or that the method of anonymising was not proportionate prior to the security incident.

f) Aggregate data, reference populations and linkage

- Para 17 is too broad if read to mean that aggregate data relates to individuals whenever any de-aggregation technique exists that could potentially single out a unique value. We would like to request a clarification that the technique must yield information about an identifiable person through means reasonably likely to be used e.g. real access to relevant secondary data sources, not just the ability to single out.

- The No Linkage analysis should account for the reference population, equivalence classes, population size, and the availability of the exact matching attributes. Quasi-identifiers should not be treated as absolute identifiers.

- The apparent tension between para 17 and para 58 should be resolved. Aggregated data is of the utmost importance to health research, e.g. for publications and presentation of results. The Guidelines should therefore not create doubt that correctly aggregated data can be anonymous data.

g) Mixed datasets, rare records and recognition of established methods

- Treating an entire dataset as personal data when one unique data record potentially remains identifiable may create significant consequences for rare-disease, registry and real-world data. The Guidelines should explain how to handle such datasets in practice and the options for continued use of the other data in the dataset.

- We request that the EDPB recognise established scientific and industry practice, including statistical disclosure control methods and standards such as ISO/IEC 27559, as supporting evidence, while remaining technology neutral. For example, risk-based anonymisation in the life sciences industry commonly applies a maximum re-identification risk threshold of 0.09, as used under EMA Policy 0070. This probability threshold roughly corresponds to k = 11 under the k-anonymity model. Rather than prescribing specific thresholds, the Guidelines should confirm that such established methods can be relied on, leaving organisations room to set thresholds suited to their context.

- The Guidelines would benefit from incorporating the existing concepts of rule-based and risk-based anonymisation and describing how these link to the contextual approach and the simplified approach (paras 47-48). In particular, the Guidelines should confirm that established risk-based methods can be used to carry out the contextual assessment. It

would also be beneficial to provide examples of how risk-based anonymisation is currently used, for example under EMA Policy 0070 and in scientific publications. Otherwise, there is a risk that the Guidelines will displace quantitative methods already in use within the life sciences industry and lead to increasing inconsistency between assessors.

h) Reassessment, historical datasets and operational scalability

- The Guidelines point out that reassessment is good practice and that risk increases over time but do not specify triggers or frequency. Reassessment should be proportionate and triggered by material changes capable of affecting the original conclusion, rather than by a universal fixed periodic review cycle or continuous monitoring obligation. An earlier documented assessment should remain valid unless a material development undermines one of its assumptions. This would provide practical certainty and predictability without insulating outdated assessments from review.

  Material triggers should include relevant changes in access, actual incidents, public or commercial comparison data, intended use or available identification techniques.

  Technological development should affect an existing assessment where it is concrete and evidenced, for example because a suitable tool has become readily available and performs reliably on comparable data. Generic predictions that AI may improve should not overturn an existing assessment. The Guidelines should also recognise that an AI-generated match may be inaccurate and require consideration of validation, confidence and false-match rates. The Guidelines should also explain the proportionate response where previously published data cannot realistically be recalled.

- We request a risk-based lifecycle approach also taking into consideration the potential harm re-identification would have on the data subjects and how this could diminish over time. For example knowing that someone is pregnant during their pregnancy versus 5 years after the child was born does not have the same potential impact on that person. As such the choice of safeguards and review effort should consider how the potential consequences of re-identification may change as data age. This should inform proportionate risk management without replacing the assessment of whether a person is identifiable.

- Organisations processing many similar datasets (e.g. clinical trial data) need to be able to reuse validated standard frameworks at the level of a processing activity, documented reasoning, data classes and repeatable controls where assumptions remain applicable, with case-by-case specific checks only needed for material differences. A fully bespoke assessment for every single dataset would create substantial burden, impede secondary research use and stifle innovation.

Specific Proposals

Reference	Issue	Proposed consultation comment
2.2, para 12	Meaning of entity and intra-group use	Please clarify whether "entity" refers only to separate legal entity or can include operationally separated functions within the same legal entity. The Guidelines should explain when data may be anonymous for a department or function that lacks access to identifying information, despite another function within the same company retaining such information, provided effective separation and controls are in place, and the limits of such an assessment.
2.2, Example 3	Processor perspective	We agree that Example 3 correctly applies the controller's perspective to the processor. The controller determines the purposes and means of the processing and is therefore the operative entity for the assessment. Please ground this conclusion more explicitly in Articles 28 and 29 GDPR: the processor acts on documented instructions and under the controller's authority, so identification means attributable to that controller cannot be disregarded within that relationship. Distinguish this from a recipient determining its own purposes, whose own controller perspective should be assessed contextually.
2.3, para 17; 3.4.1, para 58; 3.4.3, Examples 16 and 17	Aggregate data	Please clarify that aggregate data relates to an individual only where a technique can realistically derive information about an identified or identifiable person using means reasonably likely to be used. The Guidelines should describe relevant scenarios, including differencing, overlapping aggregates and reconstruction, and reference suitable technical methods. Please identify the techniques relied upon in para 17 and explain the circumstances in which each is realistically likely to be used, rather than treating their mere existence as sufficient. Please reconcile para 58, under which properly aggregated data usually satisfies No Record Isolation, with para 17 on techniques that can reveal information from aggregates. The distinction between group-level information, reconstruction of individual records and specific meaningful inference should be made explicit. Explain the limits of Examples 16 and 17, including their small populations and overlapping groups, so they do not imply a general weakness in aggregation. Show how group size, query design, suppression and repeated outputs affect reconstruction risk. Add pharmaceutical research, registry and research-collaboration examples distinguishing generalisable population-level knowledge from specific, meaningful inference about an identifiable person.
2.4, para 19; 2.4.1, para 25 and Example 6	Identification beyond uniqueness	We suggest clarifying that distinguishing an individual from others requires not only a unique attribute combination but also a realistic means to connect that combination to the individual and treat that person differently. A transient unique measurement should not by itself establish identifiability where repeat measurement or matching is not reasonably likely. Extend this explanation to device identifiers and browser fingerprints, including Example 6: distinguishing a record, device or session does not by itself establish a reasonably likely connection to a natural person.
2.4, para 20	Alternative access routes	The assessment of means reasonably likely to be used should take account of whether the same identifying information could be obtained through a

		materially easier route. This factor can help assess whether a complex linkage or intrusion scenario is realistic and proportionate rather than merely conceivable.
2.4, para 21	Chains of entities	Please clarify the limiting principle for chains of transfers and third-party capabilities. The current wording may classify data as personal where identification depends on a remote combination of entities and means. The assessment should require a realistic connection, likelihood of access and plausible combination of the relevant means. Align para 21 with para 28 by assessing the likelihood of each step and of the complete scenario, including receipt of data, access to additional information, cooperation or disclosure, accurate combination and connection to a natural person. A conceivable step should not determine the outcome where the whole pathway remains unrealistic.
2.4.1, Example 5	CCTV example	Please revise the example to distinguish isolation within a video frame from identification of a natural person. A suit and a dog may single out one visible record, but additional means, such as a visible face, persistent tracking or matchable external information, would ordinarily be needed to identify the person.
2.4.2, para 26	Safeguards and impossibility/disproportionate effort	Please confirm that appropriately implemented technical, organisational and contractual safeguards are relevant when assessing whether identification is impossible or would require disproportionate effort, while recognising that safeguards must be effective in the circumstances. Please restore the qualification practically impossible in the Breyer formulation discussed in para 26, rather than implying that absolute impossibility must be demonstrated, alongside the alternative that the relevant means are prohibited by law.
2.4.2, paras 26, 28, 29	Capable actors and proportionality	Please provide a structured proportionality test covering the number and type of capable actors, their access, the expected benefit, cost, effort, legal and contractual consequences, organisational and technical safeguards and the availability of easier alternative routes. This would make the "means reasonably likely to be used" test more operational.
2.4.2, Example 7 / para 34	Contractual controls	The Guidelines should explain when reliable, verifiable and enforceable contractual controls may materially reduce the likelihood of re-identification. Contractual safeguards should be given weight according to how reliable and enforceable they are in practice. A contractual prohibition should not be equated with a prohibition by law, but re-identification bans, audit rights, access restrictions and onward-transfer controls can be meaningful mitigating factors when combined with technical and organisational measures. Contractual safeguards should not be disregarded merely because amendment or breach is conceivable; assess the likelihood of those events. Include permitted-use limits, and explain why the re-identification prohibition in Example 7 is insufficient in its particular factual setting. Reliability, verifiability and enforceability should not be interpreted as requiring routine audits in every case regardless of risk, including agreements with contract research organisations, registries and academic partners.
2.4.2, para 29(a)	Time-varying attributes	Please add that the properties of data include whether attribute values are stable or change over time. Physiological, behavioural or cognitive measurements may have limited value for possible re-identification where repeat values are unlikely to remain sufficiently similar.
2.4.2, para 29	Easier sources and confidentiality	Please add whether identifying information can be obtained more easily from another source. The Guidelines should also address circumstances where professional confidentiality or another effective legal restriction prevents access or re-identification, rather than considering only laws that enable access. Where disclosure or re-identification is prohibited by law, this should be taken into account in the Recital 26 assessment of means reasonably likely to be used. Please clarify, in light of Recital 26 and Breyer, how

		identification through lawful, accessible means and public comparison material should be distinguished from routes dependent on cyber intrusion, stolen information or another unlawful act. The latter should principally be treated as security and breach risks. Explain how this distinction applies to concrete, evidenced unlawful-access scenarios so that security planning is not equated with the legal identifiability test.
2.4.2, para 30	Rogue employees and related persons	Please explain when rogue employees, partners, friends or other related persons must be included as relevant entities. The assessment should consider whether safeguards are effective, whether the data reveals new information e.g. potential consequences of re-identification, and whether realistic access and concrete evidence justify their inclusion, without assuming that effective safeguards will fail.
2.4.2, para 30	Law-enforcement and malicious actors	The list of possible actors should be applied proportionately. Where an actor can obtain the source information directly through a simpler lawful or unlawful route, the likelihood of undertaking a complex re-identification exercise against anonymised data is less likely.
2.4.2, Example 8	Evidence for interception scenarios	Please make the example more explicit about the evidence supporting likely interception and compromise. A purely hypothetical possibility should not by itself trigger an indefinite cycle of reassessment. The example should distinguish a concrete risk, supported by vulnerabilities or comparable incidents, from general speculation. We recommend distinguishing the identifiability test from security planning. An inability to guarantee that an incident will never occur should not itself defeat anonymity. A later successful attack may require containment, assessment and reassessment, without necessarily proving that the dataset or model was personal data before the event.
2.4.2, para 35	Predictable review triggers	Please specify material review triggers, distinguishing good practice from mandatory obligations. Reassessment should be linked to relevant changes in access, actual incidents, public or commercial comparison data, intended use or available identification techniques. Generic predictions about AI or technical progress should not by themselves require reassessment of every historical dataset, and the Guidelines should explain the response where published data cannot realistically be recalled. Later accidental identification should prompt an appropriate response and reassessment, without automatically invalidating an earlier assessment based on the means reasonably foreseeable at the time; an event with an insignificant likelihood can still occur.
2.5.1, para 36	Mixed datasets and rare records	Identification of one record should not automatically determine the status of the remaining records. Assess whether it materially increases their re-identification risk and explain when proportionate controls allow the remaining data to continue to be treated as anonymous. Please provide practical guidance on separating and handling anonymous and non-anonymous portions of a dataset, especially for rare-disease, registry and real-world data. A rare or unusual record should prompt an identification assessment, not automatic removal of clinically important information. The implications of one residual edge case for the entire dataset should be explained alongside proportionate targeted measures such as suppression, generalisation, access restriction, output checking or separate processing.
2.5.2, para 38	Research legal basis	Please explain how the requirement for a legal basis for anonymisation corresponds with GDPR Article 89, especially for circumstances where anonymisation is used as a safeguard for scientific research. The Guidelines should include a research example confirming that the existing Article 6 basis and Article 9(2) condition may be applied where anonymisation forms part of the same activity and purpose. The Guidelines should also include a description of the requirements for preparation of research data for a different research

		purpose and if this requires further assessment. The text should also clarify that Article 89 safeguards do not require a standalone legal basis. Please confirm that the presumption for the existing Article 6 basis and Article 9 condition also covers compatible further scientific research and is not restricted to identical purposes. Anonymisation undertaken as a safeguard within lawful processing should not automatically be treated as an independent purpose. Researchers should not face greater uncertainty when anonymising data than when continuing lawful research with identifiable data.
3.1, para 44	Reusable frameworks	Please clarify that organisations may implement validated standard assessment frameworks for recurring categories of data and processing (e.g. risk-based anonymisation applying a maximum re-identification risk threshold, such as 0.09 under EMA Policy 0070), with documented case-by-case specific checks and deviations. A case-by-case specific assessment should be able to reuse validated methods and documented reasoning where the relevant assumptions remain applicable, subject to checks for material differences including population size, unusual events, recipient access, data environment and intended use. Requiring a wholly bespoke assessment for every similar dataset would create substantial operational burden and may impede responsible secondary use thus limiting research and innovation.
3.4.1, para 55	Record isolation and time	Please clarify that a unique record does not on its own establish identifiability, and that the stability of unique values over time is relevant. The later singling-out analysis in paras 98-100 should be linked here to reduce the risk of para 55 being applied as a standalone rule.
3.4.2, paras 59-61	Scope of linkage analysis	The No Linkage test should be bounded by the relevant reference population, the availability and reliability of matching attributes, and the perspectives of entities that can realistically access both datasets. Requiring knowledge of all public and commercial data across jurisdictions would be impracticable. The mere possibility that a record or attribute supports an inference should not determine the outcome, as linkage does not always amount to identification. Explain expressly how effective technical, organisational and contractual controls limiting access to auxiliary information affect the No Linkage assessment.
3.4.2, paras 59-65	Equivalence classes	Please acknowledge that population size, equivalence classes and small-cell analysis are relevant evidence. A postcode or demographic combination cannot be evaluated uniquely; its linkage value depends on the population represented, attribute resolution and availability of auxiliary information.
3.5.1, paras 85-87; 3.4.2, Example 11	Statistical evidence and legal conclusion	Please add guidance explaining how existing statistical evidence supports the legal conclusion that identification is not more than insignificantly likely. The assessment should identify the relevant reference population, account for known study participation, false matches and confidence, and distinguish attempts to identify a specific individual from attempts to identify anyone in a dataset. Apply para 87's accuracy and confidence standard to Example 11. Test basic consistency, including date, geography and matching accuracy, and account for the timing of comparison information. An apparent match that does not meet that standard should not be treated as identification.
3.5.2, paras 89-94, Example 22	Controlled access versus public release	Please distinguish public release from controlled access by approved researchers in a restricted environment. Example 22 treats leakage as a precautionary assumption and should explain what evidence would justify that assumption, e.g. why human error or a hacking attempt should be presumed to succeed despite state-of-the-art cybersecurity and technical and organisational controls. The assessment should consider realistic access and the capabilities of authorised users rather than assuming that all controls will fail.

		The mode of disclosure should determine the relevant population of capable actors and the scope of supporting evidence; a threat model for unrestricted publication should not automatically be applied to named, contractually bound recipients.
3.5.2, paras 92-93	AI and changing capability	The discussion of AI should be accompanied by a proportionate temporal framework. Please clarify how foreseeable developments should be evidenced, how safety margins should be set, and when future capability becomes sufficiently concrete to affect a present assessment. Without this clarification it remains a hypothetical consideration that creates uncertainty in how to interpret the Guidelines. Please also explain when a technique is sufficiently accessible, accurate and affordable to be considered reasonably likely to be used. A concrete trigger could be an identification tool becoming readily available and performing reliably on comparable data, with accessible comparison data for the relevant population. The ability to run an algorithm or generate a possible match is not sufficient. Require validation of AI-generated matches, including confidence and false-match rates, before treating them as evidence of identification.

Concluding recommendation

Novo recommends that the final Guidelines confirm four principles that keep the contextual approach workable for scientific research:

- Data should not be treated as personal data without a concrete identification scenario.
- Identification requires a link to a natural person, not merely the isolation of a record.
- A single record should not determine the status of the whole dataset.
- The relevant means should principally be lawful and be assessed in the actual disclosure setting.