Sikt

European Data Protection Board (EDPB)	**Our date** **30.10.2026**	**Your date**
	Our reference **26/00949-5**	**Your reference**
	Ugradert	**Executive officer** **Lasse Raa**

Sikt's comments on EDPB Guidelines 02/2026 on Anonymisation

Sikt – Norwegian Agency for Shared Services in Education and Research welcomes the opportunity to submit comments on the European Data Protection Board's Guidelines 02/2026 on Anonymisation (hereinafter "the Guidelines"), adopted on 07 July 2026 and submitted for public consultation.

Sikt is a Norwegian agency under the Ministry of Education and Research providing services, infrastructure, and expertise to the Norwegian higher education and research sector. In this capacity, Sikt supports a wide range of scientific research activities across universities, university colleges, and research institutes. Among other things, Sikt acts as resource centre, assisting institutions with compliance under the GDPR. Each year, Sikt assesses GDPR compliance in approximately 12,000 new research projects on behalf of our customers. Activities range from the bachelor level to large scale national and international research projects across all fields. Sikt also provides Data Protection Officer services to more than 20 institutions in higher education and research. Furthermore, Sikt is mandated as a national research data infrastructure, managing data on behalf of the research community and providing digital solutions for storage, sharing, merging and analysing various types of data.

In this context and based on our experience, Sikt wants to offer comments and observations on several specific aspects of the Guidelines where we believe further clarification and practical guidance would be beneficial.

1. Datasets containing a mix of anonymous and personal data

Para. 36 states that "the dataset as a whole should only be considered anonymous if the anonymisation is effective for all of the included individuals" and that if a dataset "contains a mix of personal and anonymous data, the entire dataset should be treated as containing personal data (and therefore within the scope of the GDPR) in any situation where the separate parts of the dataset are not treated separately".

The threshold for separate treatment is unclear. We ask the EDPB to clarify this, in particular whether it would, for instance, be sufficient that the controller knows which individuals in the dataset are anonymous and which are not.

It is also unclear what the practical implications are if data relating to anonymous individuals should be considered to fall within scope. Data subject rights, information obligations and the requirement of a legal basis cannot meaningfully apply to data that cannot be linked to any individual.

We ask the EDPB to clarify whether the point is mainly about information security, meaning the whole dataset should be protected at the level required for personal data. If so, the Guidelines should say this explicitly.

2. The No Inference criterion

The definition of a "specific and meaningful" inference (para. 67) leaves many questions open. The line between general or population-level knowledge and an inference about an individual is hard to draw. Read broadly, the criterion could require controllers to anticipate every possible inference, including

those drawn from AI models or synthetic data. That is hard to reconcile with a risk-based test that only requires the risk to be insignificant. More concrete examples and thresholds would help.

3. Information regarding the anonymisation process

Para. 40 states that controllers should comply with their transparency obligations under Articles 5(1)(a) and 12–15 GDPR regarding the processing relating to anonymisation. Among other things, they should state clearly that personal data will be processed in order to produce anonymous data. Sikt agrees that data subjects should not be misled about whether their data are anonymous. However, the practical consequences for scientific research are unclear.

In research, anonymisation is often a routine step and a safeguard under Article 89(1), for example at the end of a project, before archiving or before data is shared. It would be helpful if the EDPB could clarify:

- whether general information at the time of collection is sufficient, for example that data will be anonymised or deleted at the end of the project, or whether more specific information about the anonymisation process or purposes for use of the anonymised data is expected;
- whether anonymisation preceding further use of anonymised data that was not foreseen in the initial information is considered to be a change to the processing that requires updated information, and if so, how this relates to section 5.5 of Guidelines 1/2026 on changes to processing operations;
- how the obligation applies to existing research datasets.

Sikt raised related questions in its comments on Guidelines 1/2026, regarding derived variables (para. 91) and changes to processing (section 5.5). We encourage the EDPB to make sure the two sets of guidelines are consistent on this point, and to include a worked example of anonymisation in a research context.

4. Qualitative data and unstructured information

The framework and almost all examples in the Guidelines assume structured, record-level or aggregate data in tables (e.g. Examples 10, 12 and 24). In scientific research, a large share of the data that are anonymised for archiving, reuse or publication is qualitative: interview transcripts, field notes, open-ended survey answers, audio and video. In such data, identification seldom happens through a unique combination of variables, but rather through context, such as a described event, a workplace, a way of speaking, or details about family members and colleagues. Para. 30 rightly points out that an individual's partners, friends and colleagues may be relevant entities. In qualitative research these are often exactly the people able to recognise a participant, or a third person mentioned in an interview, from contextual details.

It is not clear how the three criteria should be applied to such data. In particular, how No Linkage should be assessed when the possible linking information is someone's personal knowledge rather than another dataset, and what level of residual recognisability among people close to the participant is acceptable. We would welcome guidance and at least one worked example on anonymising qualitative research data, for instance an interview transcript intended for archiving or for quotation in a publication.

5. Data that become personal data again over time

Para. 35 states that if the likelihood of identification later becomes more than insignificant, previously anonymous data should again be considered personal data and the controller will be accountable for any processing of it. Para. 88 recommends adequate safety margins. We recognise that re-identification risk changes over time, especially with the developments in AI described in paras. 92–93. However, research data and results are often published openly or deposited in archives. Once published and downloaded by external parties, they cannot in practice be withdrawn, and the original controller has no control over third parties' further use of the original dataset.

We would therefore welcome clarification of what accountability under para. 35 entails for data which have been lawfully published or shared as anonymous on the basis of a careful, documented assessment. This includes which entity is accountable when the data are processed by others, what is

expected of the original controller in terms of reassessment, and whether a well-documented assessment at the time of publication is sufficient to show compliance. More concrete guidance on what an adequate safety margin looks like for openly published research data would also help controllers decide between open publication and controlled access.

6. Thresholds for identifiability

The Guidelines appear to use two different thresholds for identifiability. Para. 22 states, in line with EDPS v SRB, that the likelihood of identification does not need to be zero but should be "insignificant in reality". Para. 21 states that data should be considered personal, both for the transfer and for the transferring entity, if a transfer to a third party is reasonably likely and "it cannot be ruled out" that the recipient has means to identify the individual. Read literally, "cannot be ruled out" comes close to requiring zero risk, which seems hard to reconcile with para. 22 and with the risk-based approach of the rest of the Guidelines.

7. Proportionality of the assessment

The Guidelines explain what must be assessed, but not what effort is sufficient for a defensible conclusion. The No Linkage criterion and the assessment of means reasonably likely to be used could in principle be pursued indefinitely. For smaller controllers and research environments, this creates an incentive either to over-invest or to treat data as personal by default.

We ask the EDPB to clarify:

- that the depth of the assessment and its documentation may be scaled to the risk of the release, including the sensitivity of the data and whether access is open or controlled;
- what constitutes a reasonable search for linkable datasets;
- whether one assessment may cover recurring releases of the same type, and whether recipients may rely on an assessment documented by the disclosing entity; and
- the minimum content of the documentation, preferably as a template.

8. Perspective and roles

Under para. 15, data processed on behalf of a controller are assessed from the controller's perspective, whereas independent recipients are assessed from their own perspective.

We ask the EDPB to address, with examples:

- providers that process data both on behalf of the controller and for their own purposes; and
- whether information held by a processor, such as account data or IP addresses, counts among the means reasonably likely to be used by the controller.

With best regards

Vigdis Namtvedt Kvalheim
Director, Research Services Division

This document is electronically approved and therefore has no handwritten signatures.

Recipients:
European Data Protection Board (EDPB)