Statement on the Public Consultation: Guidelines 02/2026 on Anonymisation

The draft Guidelines make an important and valuable contribution by modernising the concept of anonymisation, especially in light of recent developments in data analytics, large-scale data linkage, and artificial intelligence. In particular, the recognition that anonymisation must be assessed contextually rather than through a simple checklist is welcome. The emphasis on linkage, singling-out, and inference risks reflects the realities of contemporary data processing and should improve the quality of anonymisation assessments across sectors.

Although I appreciate the fact that the guidelines are broadly consistent with the GDPR framework for scientific research, further clarification would be beneficial regarding the relationship between anonymisation and the scientific research framework, even when taking into account the Guidelines 1/2026 on processing of personal data for scientific research purposes.

The draft Guidelines establish a demanding standard for concluding that data can be regarded as anonymous. In practice, many research datasets, particularly in health research, longitudinal studies, social sciences, humanities research, and research archives, are likely to remain personal data even after extensive de-identification measures have been applied. This is not inherently problematic, as the GDPR expressly permits the processing, retention, sharing, and secondary use of personal data for scientific research under appropriate conditions. However, the interaction between these two frameworks could be articulated more clearly.

Without additional clarification, research organisations may face uncertainty as to whether anonymisation is expected before research data can be archived, shared through research infrastructures, or reused for future scientific purposes. There is a risk that the Guidelines could be interpreted as implicitly favouring anonymisation in circumstances where the GDPR instead contemplates the continued use of pseudonymised personal data protected by Article 89 safeguards.

To ensure legal certainty and consistent application across the EU in the context of scientific research, I would suggest that the final Guidelines more clearly explain:

- how controllers should assess whether anonymisation would impair the scientific validity, utility, reproducibility, or long-term value of research data;

- the circumstances in which pseudonymisation and other Article 89 safeguards may provide an appropriate alternative to anonymisation, especially in scientific research contexts.
- the extent to which organisational and governance safeguards may be taken into account when assessing whether particular means of identification are reasonably likely to be used in practice; and
- how the recipient-specific assessment of identifiability should be applied in the context of research infrastructures, repositories, archives, and international scientific collaborations.

Additional examples addressing controlled-access research repositories, secondary use for research, qualitative research archives, and long-term research infrastructures would be particularly valuable.

Such clarifications would strengthen legal certainty regarding the interaction between anonymisation and the safeguards applicable to scientific research under Article 89 GDPR. This would assist controllers in determining when anonymisation is appropriate and when continued processing of personal data for scientific research purposes may lawfully rely on alternative safeguards. Greater clarity would support consistent application of the GDPR while facilitating responsible scientific research, research reproducibility, and the lawful sharing and reuse of research data within the EU.

There is also a need for further guidance regarding aggregate data. The anonymisation guidelines appear primarily focused on the release of datasets, whereas research institutions often release only aggregate results. For many universities, the most common disclosure to the public is not individual-level data but statistical outputs. Consequently, practical guidance on when aggregated research outputs can be considered anonymous may be at least as important as guidance on anonymisation of datasets themselves.