Clusit welcomes the opportunity to contribute with our comments on the “Guidelines 02/2026 on Anonymisation”.

The aim of the Guidelines is to "clarify the notion of anonymous data and provide a practical framework for determining whether data has been successfully anonymised. This guidance will therefore help to ensure that data is safely anonymised where possible and proportionate, allowing for free and fruitful use of data while preserving the rights and protections of natural persons". The Guidelines provide an extensive discussion of the concept of anonymous data and the various factors that must be taken into account when assessing whether data is anonymous. They also address, in line with the cited judgment of the Court of Justice, the question of the perspective from which data may or may not be considered anonymous, or may instead be pseudonymised.

However, the Guidelines do not appear sufficiently effective in providing the tools needed to assess in practice whether data can be considered anonymous. Moreover, they do not sufficiently distinguish between cases in which data is genuinely anonymous and therefore falls outside the protections of the GDPR, and cases in which data is anonymous from the perspective of one entity but remains personal data from the perspective, for example, of the controller, which must therefore continue to ensure its protection. In some respects, the discussion is rather discursive: it repeatedly addresses the relevance of a particular factor in supporting the anonymisation of personal data, only to revisit it later as a factor weighing against anonymisation, generally without providing criteria that could genuinely guide the balancing of these considerations in an assessment. Nor do the examples provide a sufficiently clear basis for drawing the distinction.

It should also be borne in mind that the Guidelines will, it is hoped, be used by a wide variety of data controllers. Although an understanding of the subject matter is indispensable, the Guidelines should also provide effective guidance for assessments carried out, for example, by a small-business controller, with the limited resources and expertise it can devote to the assessment. By way of illustration, and in keeping with Clusit’s area of expertise, consider the unauthorised access to data by cybercriminals, who may subsequently make the data public. In principle, anyone possessing additional information that could be used to enrich that data and re-identify the data subjects would then be able to do so. This issue is addressed at least in the following passages:

30 *The GDPR does not lay down any conditions regarding the entities who are able to identify the individual. Relevant entities may, therefore, include anybody directly or indirectly receiving the data, including unauthorised and malicious actors. Depending on the circumstances at hand, these may include (among others) 25:*

...

Cybercriminals. These entities may simply break the law to gain access to data not available from a law-abiding entities’ perspective.

32 *The CJEU has stated that means do not need to be considered if, in reality, the likelihood of their use is insignificant because they are prohibited by law28. It should be noted that these cases involved actors who were properly limited by the rule of law, so that the law effectively prevented the use of the means in question. Indeed, we can generally assume that people will follow the law. However, there may also be cases where illegality does not provide any meaningful barrier to identification, because the entities involved will not actually be restricted by this prohibition.*

33 *The general assumption that people will follow the law can, therefore, be rebutted if there is sufficient evidence of a concrete risk that unlawful means are nevertheless reasonably likely to be used. The assessment of this risk should consider, in particular, evidence that the prohibition in question is not respected by the entities involved. Relevant factors may include, but are not limited to:*

- evidence that the legal prohibition is not effectively monitored, enforced and sufficiently dissuasive;

- evidence that the gains from breaking the legal prohibition might outweigh costs, effort and potential risks of doing so;

- evidence of a situation where data relevant for re-identification (i.e. both the given and, potentially, the additional data useful for re-identification) is especially vulnerable to illegal access using the prohibited means, due to a lack of protection; and

- evidence that the prohibition has previously been breached or circumvented in comparable situations.

Example 8: *...However, the media company also recognises that there is a risk of the shared data being unlawfully intercepted by unintended third parties. These third parties include entities located outside the scope of applicable Union and Member State law, not effectively bound by such laws prohibiting unauthorised access. Moreover, there have been several incidents of illegal access affecting this and other media company's IT systems storing personal data. Since the media company's out-of-date security measures mean that those third parties could also likely hack into the media company's servers, these unintended receivers of the data would be able to access the privileged information and therefore re-identify the individuals through means reasonably likely to be used. The shared information would not, therefore, be considered anonymous from their perspective. The media company therefore decides that it will upgrade its network security before re-testing the anonymity of the data.*

42 *In some cases, a security incident may lead to a reassessment of anonymity. This is particularly the case if the assessment of anonymity relied on certain information being kept confidential, but the incident means that the individuals can now be identified with means reasonably likely to be used. Where this is the case, the entity which suffered the incident should then determine its role with respect to the processing and whether it needs to fulfil any obligations under Article 33 and 34 GDPR*

88 *When using the contextual approach, it is important to recognise that the capabilities of different entities may vary over time. For example, a data leak could provide certain entities with access to new datasets which were previously unavailable to them. This would, in turn, allow those entities to link information with means which were, at least from that entity's perspective, previously unavailable but are now reasonably likely to be used, leading to a violation of the No Linkage criterion. To ensure the proper protection of data subjects, the EDPB therefore recommends that analyses carried out under the contextual approach should incorporate, as much as possible, adequate safety margins ensuring that the assessment remains robust as the entities' capabilities evolve*

89 *If the data is subject to appropriate security controls (particularly regarding the confidentiality of the data), access to the data will constitute a mean reasonably likely to be used only for the authorised users (and not for unauthorised actors).*

90 *Limits to access should be assessed holistically, taking into account the capabilities of the different entities and any contextual elements which may have a greater or lesser effect on the means they are likely to use. Technical, organisational and contractual measures can help to restrict access, but are not by themselves sufficient to make data anonymous;*

91 *Appropriately implemented encryption may be used to limit access to data (or, at least, limits access to a form of the data which is intelligible) and, in this way, can help to support successful anonymisation. Nevertheless, encryption is, by its very nature, intended to be reversible and differs from anonymisation. It is, therefore, important to assess if entities with access to the data could still decrypt it with means reasonably likely to be used, even if they do not currently have the decryption key (e.g. through cryptanalysis).*

Example 22 *...The table has not been publicly released, and the research institute currently employs strong state-of-the-art cybersecurity measures, as well as other technical means to restrict unauthorised access to this dataset and prevent its members from creating any copies. Out of caution, the research institute considers that, especially over a long period of time, it is possible that the data will nevertheless leak. This could include, among other things, leaks due to human error or hacking attempts from external actors. The research institute therefore concludes that the table is accessible, through means reasonably likely to be used, by, at least:*

...

3 individuals who might hack into the system and steal and distribute the data

As can be seen, the Guidelines indicate that security measures may be regarded as effective in preventing unauthorised access (for example, in paragraph 89). Yet Example 22 takes a precautionary approach, pointing out that even with state-of-the-art security measures, the possibility of the data becoming accessible cannot be ruled out. Although the quoted passages certainly address many aspects of unauthorised access, they still make it difficult to assess the effectiveness, or even the relevance, of security measures. A cybersecurity specialist asked to contribute to such an assessment would immediately point out that security measures, even state-of-the-art ones, are generally tools for mitigating risk rather than eliminating it. They would also note that many attacks can succeed with limited resources when the right conditions arise. Vulnerabilities disclosed recently as a result of analysis performed using AI tools are certainly an example, although this was also true before such tools became available. The specialist would therefore need criteria that are as objective as possible to assess the adequacy of security measures in preventing unauthorised access from leading to the disclosure of data and, potentially, its enrichment with additional information. Reliance on historical information about previous breaches is well known to be of limited value, given the rapid evolution of risks and the surrounding context, as well as the generally limited availability of usable historical data.

The passages from the Guidelines quoted above make it difficult to identify guidance on which a controller could reasonably rely if it assessed data as anonymised and that assessment were subsequently challenged, for example, by a supervisory authority. Given the inherently high risk faced by a controller deciding to treat data as anonymous on the basis of the revised Guidelines, it would be advisable for the Guidelines to specify sufficiently clearly the circumstances in which the controller would not be subject to sanctions, even if the assumptions underlying its assessment were subsequently proven incorrect by actual events, a scenario explicitly contemplated, for example, in paragraph 42.

In conclusion, we suggest providing more explicit criteria and decision-making guidance for assessing the different factors, for example in an appendix following the extensive discussion in the current text.