noyb - European Centre for Digital Rights
aße 172/4/3/2

Subject: ***noyb observations on Guidelines 01/2025 on Pseudonymisation (version for public consultation)***

To the European Data Protection Board,

noyb welcomes the opportunity to submit comments on the Guidelines 01/2025 on Pseudonymisation ('*the Guidelines*').

1. Introduction

noyb welcomes the recent Guidelines and congratulates the authors on the very detailed and useful guidance. In the interest of efficiency, our feedback is consequently only very short and focuses mainly on the elements that may benefit from further input.

Nevertheless, we would like to express our support for the general approach taken in the document. For instance, we agree with the definition of a pseudonymisation domain (albeit with some remarks, see below § 2.1). We find it useful to reiterate that pseudonymous data is personal data. And that pseudonymous data remains "personal" even if the additional information is deleted, if the conditions for complete anonymity are not met [§22]. The clarifications about the pseudonymization domain are clear and useful (§ 2.2). Likewise, we support the EDPB's position that, in certain contexts, the value of pseudoanonymised data can be used for the authentication of the data subject and the exercise of GDPR rights [§§ 77-79].

Beside *noyb*'s general agreement with the Guidelines, we would like to comment on some elements.

2. Additional Feedback

noyb identified the following cases in which the guidelines could be improved:

2.1. Remarks on pseudonymisation domain

Pseudonymisation domain as we understand the concept defined in the guidelines is: the people to whom the pseudonymised data is available, but who are lacking the means to re-attribute the pseudonymised data to the data subject [§35]. *A contrario* people outside of the pseudonymisation domain would thus possess both the pseudonymised data and means of reattribution.

It would therefore most likely be easier for a controller to delineate who is outside of the pseudonymisation domain than who is in it. After all, only the people who have access to both the pseudonymised data and the means of re-attributing the data to the data subject are outside of the domain, while e.g. a hacker or accidental recipient would typically be in the domain.

Since the obligation of the controller under article 5(2) is to show adherence with the principles, and since it's only possible to prove a positive, it follows that a list of people who have access to data is better proof for adherence to the principle of integrity and confidentiality [Art 5(1)(f)] than a list of people who are assumed not to have that access.

We therefore suggest a redefining of pseudonymisation domain, to instead mean: the people to whom pseudonymised data and the means to re-attribute that data to an identified data subject is available. This way controllers can more easily assign and control the size of the pseudonymisation domain, and use the concept as a means to show adherence to the principles of the GDPR.

Our suggestion is thus to change § 35 to something akin to the following: "*Controllers **should** define the context in which pseudonymisation is to **include** attribution of data to specific data subjects, generally on the basis of a risk analysis. They subject the additional information to technical and organisational measures to ensure that the pseudonymised data cannot be attributed to data subjects by persons operating **outside** that context. This means in particular that additional information that would enable attribution is kept **within it**. These guidelines call this context (with the people operating in it and its attending physical and organisational aspects, including the IT assets available) the pseudonymisation domain.*"

2.2. Clarify the requirements of pseudonymization domain

Pseudonymisation can reduce the risks associated with data processing for example in the case of accidental loss or hacking.. The Guidelines provide extensive guidance on the requirements of effective pseudonymization. However practice shows that controllers **largely over-inflate pseudonymization claims and have seriously underdeveloped protections** when processing such personal data. Clear red lines are therefore required.

To begin, the additional information needed to re-identify the data subject should only be available to persons "*specifically **authorised** for this purpose*" [§32].

Pseudonymised personal data should only be shared with a **pre-defined** set of **recipients** [§ 48]. When this happens, controllers should also put in place measures to ensure that actors within the pseudonymisation domain are not able to reverse the pseudonymisation [§ 40]. The additional information **must not** be available to such third parties [§ 40].[1] In any case, the easier it is for third parties to find 'additional information', the weaker the pseudonymisation will be [§21].

For effective pseudonymisation within a single organisational unit or a set of legitimate recipients, all involved controllers and processors should choose appropriate technical and organisational means-possibly including legal safeguards (e.g., contracts) but only if these can be **effectively enforced** [§ 39]. When one or more independent controllers are included in the domain, **more extensive measures** must be adopted [§51].

All intended recipients of the pseudonymised data need to "*demonstrably assure*" that the pseudonymised data **are not disclosed** to unauthorised recipients beyond the defined domain [§ 51]. It should be unreasonably difficult for these third parties to find additional information for singling out the data subject [§60]. It is 'good practice' that the receiving controller informs the sending controller about the processing risks on the receiving end [§72].

[1] This is the so-called 'pseudonymisation domain', a domain of parties operating on pseudonymised data. This domain can also include third parties which were not meant to receive the information but can potentially obtain it (leaks, cyber-attacks, unauthorised sharing) [§§ 35-37].

Finally, under Article 11(2) GDPR, the identity of the data subject is the pseudonymous value itself. The controller should help the data subject retrieve that value, and how it can be used to demonstrate their identity, when needed [§§ 77-79].

Currently, all this information is scattered throughout the document and not easily retrievable. *noyb* suggest inserting a new paragraph (or graph) describing the **essential requirements** of an effective pseudonymisation domain, including, pre-defined number of recipients, contractual safeguards, protections under Art. 28 GDPR, and when they are concretely enforceable. This would be a quick fix which may make the Guidelines clearer and more accessible for controllers and processors.

2.3. Insert an example on data brokers/RTB

Drawing from the previous paragraph and based on wide (ab)use in practice, we suggest providing an example describing what a pseudonimyzation domain would look like in one or more instances of data exchange between RTB/AdTech players.

This system is based entirely on the correlation of very detailed databases based on online tracking. The information collected is highly interconnected by means of cookies and other mobile identifiers,[2] which many controllers (falsely, see Recital 30) claim to be "pseudonymized" data.

The "additional information" is available to many players and in many cases "data enrichment" is not seen as wrongdoing, but understood to be an art. Furthermore, there is no clarity about the contractual safeguards adopted during the data enrichment operations as well as their concrete enforceability, as required by the EDPB. Moreover, vast information is traded and many players have access to the same datasets, which is why pseudonymization is these cases essentially pointless.[3] It is needless to say that controllers largely reject the exercise of user rights (e.g. under Article 15, 16 or 21 GDPR) based on alleged "pseudonymization".

It would be very helpful to have the EDPB's authoritative guidance on how the pseudonymization domain would look like in this specific instance, given that in our experience it is one of the most prevalent uses of Article 4(5) GDPR that data subjects are confronted with.

***

We hope these comments are useful for your work and want to congratulate the authors once again on the general approach taken in these Guidelines. We are at your disposal should you have further questions or require additional clarifications.

Vienna, 27.2.2025

Stefano Rossetti Joakim Soderberg

[2] Incidentally, let us point out that in 46 pages of guidelines, the word 'cookie' is not mentioned even once. Yet, cookies or other mobile identifiers are classic examples of pseudoanonymous data, as is also clearly mentioned in the EDPB's right of access guidelines: '*Personal data collected for behavioural advertising are usually collected by means of cookies and associated with pseudonymous random identifiers*' (Example 11, p. 25).

[3] Rocher, L., Hendrickx, J.M. & de Montjoye, YA. Estimating the success of re-identifications in incomplete datasets using generative models. Nat Commun 10, 3069 (2019). https://doi.org/10.1038/s41467-019-10933-3