Response to the EDPB on its pseudonymisation guidelines – March 2025

Google welcomes the opportunity to provide feedback to the European Data Protection Board on its *Guidelines 1/2025 on pseudonymisation*, as adopted on 16 January 2025 (the **Guidelines**). Our response contains some general observations on the Guidelines together with a number of specific points of interpretation that we would be grateful for the EDPB to consider.

At Google, we believe innovation and technological adoption stem from user and public trust. This trust is built on our ability to give our users confidence that their data is handled responsibly and in line with our market-leading security standards. It is therefore important to encourage the development and use of technologies which strengthen the security of user data whilst facilitating and promoting innovation in emerging fields. This helps drive responsible innovation, creates a safer ecosystem for internet users and can benefit organisations in a number of sectors from R&D and public health to online advertising.

Privacy protections such as pseudonymisation play an important role in facilitating responsible innovation and ensuring organisations can grow, change and develop without compromising the security of personal data. Google, like many large digital technology platforms relies on pseudonymisation as a key security measure for a number of the processing activities it carries out.

General feedback

Conflict with the Opinion of the Attorney General of the CJEU, in the EDPB vs SRB case

It is unfortunate that the Guidelines were adopted (on 16 January 2025) prior to the publication of the Opinion of Advocate General Spielman in the case of EDPS v SRB (**AG Opinion**) (on 6 February 2025). The AG Opinion directly contradicts the EDPB's view that pseudonymised data remains personal data *in all cases* when it is in the hands of a third-party recipient (without any reference to whether the receiving third-party can reasonably identify data subjects from the data). This contradiction is significant, as some of the conclusions reached, and recommendations made in the Guidelines are predicated on the EDPB's view.

We suggest that any update to, and finalisation of, the Guidelines is delayed until after the final judgment of the CJEU is published, as if the CJEU agrees with the Advocate General's view, the judgment may have far reaching consequences for organisations dealing with pseudonymised data. It is therefore important that the Guidelines are aligned with the CJEU's judgment. We would also strongly encourage the EDPB to clarify publicly that in light of the AG Opinion and the pending CJEU judgment, it will await the CJEU judgment before updating and finalising the Guidelines to take account of the judgment and feedback received via consultation. This would create a clear road map for organisations and provide certainty on the next steps in this space given the current inconsistency between the Guidelines and the AG Opinion.

Overall approach

Google welcomes the development of new guidance on this topic and fully supports pseudonymisation practices that support effective protection of personal data. In particular, we are pleased to see the EDPB's explicit acknowledgement that pseudonymisation of personal data is a technical and organisational measure, and therefore the legal basis relied on for processing personal data also applies to its subsequent pseudonymisation.

However, certain requirements and recommendations set out in the Guidelines may create unnecessary obstacles for organisations that seek to use pseudonymisation to enhance the privacy protections that are applied to personal data. For example, the Guidelines:

- state that, when assessing the means available to third-parties to re-identify the pseudonymised data, regard should also be had to the means that are reasonably likely to be used by cyber-crime actors (paragraph 42, as discussed below). This will be a challenging and potentially speculative exercise for organisations to complete in practice, as the techniques used by cyber-crime actors are constantly evolving and can be very difficult to predict. Therefore, we suggest aligning the assessment of the effectiveness of pseudonymisation with the security requirements under Article 32 of the GDPR (i.e. what is state-of-the-art *at the time* of pseudonymisation);
- set out some *absolute* criteria that must be satisfied in all cases for pseudonymisation to be considered effective, which conflicts with the rest of the Guidelines in which the EDPB is suggesting that assessing the effectiveness of pseudonymisation is a *judgment call* (based on the means *reasonably likely* to be used to re-identify the data) (paragraph 47, as discussed below). Any criteria against which the effectiveness of pseudonymisation is assessed must align to the security requirements under Article 32 of the GDPR (i.e. to take into account the state-of-the-art as *at the time* the personal data are pseudonymised);
- require that when pseudonymisation is used as a supplementary measure (to ensure compliance with Article 44 and 46(1) of the GDPR), organisations must assess the information that public authorities in the third country can be expected to process, or to be able to obtain with reasonable means (even where those means infringe legal norms in the third country) (paragraph 64 and 65, as discussed below). In practice, such an assessment would appear to require speculation regarding facts that may be unavailable to the controller and difficult to complete with a reasonable degree of accuracy; and
- require that the controller receiving pseudonymised data should re-identify the data (even where it cannot identify data subjects from it on its own) where the data subject provides it with the pseudonyms required to achieve re-identification, so that the receiving controller can respond to data subject rights requests (paragraph 78, as discussed below). This potentially undermines the reason for which the personal data was pseudonymised in the first place (e.g. to protect its confidentiality). This also conflicts with the AG Opinion which adopts the position that where the receiving controller cannot re-identify the data, then it is deemed not to be processing personal data. As such, it is not clear why the receiving controller should be required to

re-identify the data in such circumstances if provided with pseudonyms by the data subject in order to respond to a data subject rights request.

Whilst paragraphs 27-30 of the Guidelines list some limited benefits of pseudonymisation (i.e. reducing confidentiality risk, risk of function creep, the risk of data being inaccurate, and assisting in compliance with specific GDPR obligations), they could do more to promote the use of pseudonymisation by elaborating on how pseudonymisation can be integrated into GDPR compliance programs.

Pseudonymisation techniques provide a means to enhance the protection of the rights and interests of data subjects, whilst also driving broader societal and commercial benefits. These techniques support responsible innovation (e.g. by using pseudonymisation to reduce the identifiability of certain data that is used to train artificial intelligence models), create a safer ecosystem for internet users (e.g. through the adoption of PETs as part of online services) and benefit organisations in a number of sectors from R&D, life sciences and public health to online advertising (e.g. by allowing pseudonymised data to be used for research and within the adtech environment). We would suggest that this section of the Guidelines is expanded to better acknowledge and explain these practical benefits (to data subjects, businesses and society more broadly) that flow from pseudonymisation.

The implementation of pseudonymisation techniques can be resource intensive. This is especially true for more advanced techniques that involve robust technical safeguards to protect against re-identification of data, as such safeguards may require advanced technical knowledge to implement and add latency to the overall processing. Given the privacy-protective benefits of pseudonymisation, we encourage the EDPB to promote its adoption by providing more clarity about how pseudonymisation can help businesses meet their GDPR obligations. This would incentivise more widespread use of pseudonymisation across the business community by making it clear that the costs of implementing pseudonymisation are a smart investment for a GDPR compliance program.

Adoption of Guidelines in isolation from proposed EDPB guidelines on anonymisation

The Guidelines have been published in isolation from the EDPB's planned draft guidelines on anonymisation – and only a very limited number of passing references are made in the Guidelines to anonymisation.

The concepts of pseudonymisation and anonymisation are interconnected and often overlap and therefore we would have expected the two sets of guidelines to speak to each other and operate together. For example:

- both anonymisation and pseudonymisation are dependent on the same underlying concept – i.e. the *identifiability* of individuals (and the varying degrees to which this is possible); and
- there is also overlap in the criteria and factors to be applied when assessing the *effectiveness* of the anonymisation or pseudonymisation (e.g. is it reasonably likely that a third-party could re-identify the dataset through other information available to them (such as publicly available information), the state of technology and techniques available to them, the resources available to them, the costs involved in

re-identification, legal processes and avenues that may be available to obtain additional information that could be used for re-identification, etc).

Additionally, if the CJEU judgment confirms the AG Opinion, in certain circumstances, it may be possible to treat pseudonymised data in the hands of a third-party recipient as anonymised data that is not subject to GDPR. We would therefore encourage the EDPB to finalise the Guidelines in parallel to its proposed guidelines on anonymisation (as a package) in order to ensure alignment and clarity, including in relation to the *EDPS vs SRB* case.

For context, the Irish Data Protection Commissioner's guidance on anonymisation and pseudonymisation considers both these concepts in parallel, as does similar guidance from the UK Information Commissioner's Office. We hope the EDPB will consider a similar model to assist organisations in applying these concepts in practice.

No reference to AI or privacy enhancing technologies

There is no reference in the Guidelines to artificial intelligence (**AI**) and very limited recognition of the concept of privacy enhancing technologies (**PETs**). This is a notable omission given the growing prevalence of both these technologies and their rapid adoption by organisations across all sectors. In some scenarios, PETs can be utilised to supplement the protection offered by pseudonymisation, thereby further reducing the re-identification risk following pseudonymisation. This can result in the privacy risk associated with a particular activity being drastically reduced whilst simultaneously allowing for greater and more unencumbered innovation. We would recommend that the Guidelines should therefore include examples of specific PETs that could be used either as part of the pseudonymisation process, or in combination with pseudonymisation to ensure a high standard of privacy protection.

It is also crucial to acknowledge that the pseudonymisation (or anonymisation) of some data used to train AI, such as images, presents significant challenges. Determining when, for example, an image of a person can be considered pseudonymous is an open question with no universally accepted solution. The Guidelines could therefore also address the specific difficulties and nuances of applying pseudonymisation techniques to diverse data types, and encourage further research and development in this area.

Advertising

We welcome the EDPB's acknowledgment in the Guidelines that pseudonymisation has a role to play in relation to personalised advertising. At Google, we are investing in PETs and working to make them available to our adtech customers, as we believe they bring important benefits. We would suggest that further reference to these benefits should be built into the Guidelines, as suggested above in the *Overall approach* section of this response (in relation to expanding the Guidelines to acknowledge and explain the wider practical benefits of pseudonymisation).

Some good illustrations of our implementation of PETs include, (i) differential privacy, which minimises and protects personal data (by allowing for analysis of large data sets in a way where no one person's data is ever disclosed), and (ii) fully homomorphic encryption, which encrypts the underlying data in a way that Google can continue to work on it without being able to identify anyone from it. This helps Google protect the rights of individuals whilst driving responsible innovation, creates a safer ecosystem for internet users and can benefit organisations in a number of sectors.

It would be helpful for the EDPB to provide further commentary in the Guidelines on how these technologies can be used to pseudonymise personal data and help controllers meet their GDPR compliance obligations.

Assessing the effectiveness of pseudonymisation (paragraph 21, 22, 42 and 43)

The Guidelines state that in order to determine whether personal data has been effectively pseudonymised, the pseudonymising controller must assess the means that are reasonably likely to be used by that controller and others for re-identifying the data. The only example given of such means appears in paragraph 21, which refers to *"information from publicly accessible sources, such as posts in a social media or an online forum"*. The lack of further guidance on what "means" should be taken into account as part of this assessment will create uncertainty for organisations, who may struggle to carry out this assessment in the abstract. It would be helpful if further detail (including examples) was provided on the types of means that should be taken into account when making this assessment. This could include, for example, the state of technology and techniques available, the resources available, the costs involved in re-identification, and legal barriers to obtaining additional information that could be used for re-identification, etc. It would also be helpful for the EDPB to clarify that these means are assessed *prior to* the point at which pseudonymisation takes place.

For context, the UK Information Commissioner's Office updated draft guidance on anonymisation, pseudonymisation and privacy enhancing technologies (Chapter 3) lists a number of these factors for consideration when assessing whether pseudonymisation has been effectively carried out.

Paragraph 42 also states that when carrying out this assessment, controllers should consider the means that are reasonably likely to be used by cyber-crime actors to re-identify the pseudonymised data. This will be a challenging and potentially speculative exercise for organisations to complete in practice, as the techniques used by cyber-crime actors (who are motivated to commit a crime) are constantly evolving and can be very difficult to predict. Therefore, we suggest aligning the assessment of the effectiveness of pseudonymisation with the security requirements under Article 32 of the GDPR (i.e. what is state-of-the-art *at the time* of pseudonymisation).

Measuring the effectiveness of pseudonymisation (paragraph 47)

Absolute rather than reasonable criteria

Paragraph 47 appears to set out three *absolute* criteria that must be satisfied in all cases for pseudonymisation to be effective. For example, that third-parties *"are **not** able to reconstitute the original value of the attributes that have been omitted or transformed"*, or *"**cannot** link the pseudonymised data to other data relating to the same person"* (emphasis added). This conflicts with the rest of the Guidelines in which the EDPB is suggesting that assessing the effectiveness of pseudonymisation is a *judgment call* (based on the means *reasonably likely* to be used to re-identify the data).

Having such absolute criteria also creates the danger that a controller could be judged retrospectively for the effectiveness of its pseudonymisation (carried out in good faith) on the basis that the pseudonymisation could be reversed at some point in future due to technological improvements. The Guidelines must therefore ensure that any criteria against which the effectiveness of pseudonymisation is to be measured must align to the security requirements under Article 32 of the GDPR (i.e. to take into account the state-of-the-art as *at the time* the personal data are pseudonymised).

Use of pseudonymised data to single out data subjects in "other contexts"

The Guidelines state that effective pseudonymisation requires that the persons handling the pseudonymised data *"not [be] able to single out the data subjects in other contexts on the basis of what they learned from handling the pseudonymised data."* It is not clear what "other contexts" means. Reading this in line with other parts of the Guidelines, it appears that this may be referring to other contexts *within* and not *outside* of the pseudonymisation domain. It would be helpful for the Guidelines to clarify this point.

Pseudonymisation domain and available means for attribution (paragraph 20 and 35 – 43)

The Guidelines explain that the controller should define a "pseudonymisation domain" that has within it the persons with which the controller wishes to share pseudonymised data and who should not be able to re-identify it. This pseudonymisation domain could be limited to, for example, a single unit within the organisation of the controller or an external recipient. The controller must then ensure (through the use of technical and organisational measures) that *"the additional information is not to be disclosed to or used by persons processing the pseudonymised data."* (paragraph 20, 35 and 39).

This suggests that it is necessary for different people to handle (i) the pseudonymised data (i.e. those in the pseudonymisation domain), and (ii) the additional information required for re-identification (i.e. persons that are not in the pseudonymisation domain), with persons in the domain having absolutely no ability to access the additional information. Such an absolute separation may be inefficient, and in some cases, challenging to implement. For example, if a pseudonymised dataset contains quasi-identifiers that could allow for re-identification if such data is combined with publicly available data, the persons handling the pseudonymised data (i.e. those in the pseudonymisation domain) could in principle access such public data through the internet and potentially reverse the pseudonymisation. However, if the controller were to establish a small pseudonymisation domain and implement appropriate technical, contractual and/or policy measures to prohibit persons with access to the pseudonymised data from sharing such data outside of the pseudonymisation domain, or bringing external data into the domain (e.g. the publicly available data that could facilitate re-identification), such data should still be considered pseudonymous, even though the persons handling such data technically have the means to access the additional information in a different context (e.g. when acting in a different role or capacity). We would therefore suggest that the Guidelines should take into account that enforcing the separation of the pseudonymised data, and the additional information required to re-identify it, within a role or capacity-based context (through appropriate technical, organisational, contractual or policy measures) should be permitted.

It is also not clear from the Guidelines whether it is recommended in all cases to assess the effectiveness of the pseudonymisation with reference to the concept of a pseudonymisation domain. As explained above, it may not be feasible in all circumstances to implement and maintain a pseudonymisation domain with persons that have no access to the additional information. We would welcome further clarification on this point from the EDPB.

Methods of controlling publicly available information (paragraph 9 and 21)

Paragraph 9 states that pseudonymisation requires technical and organisation measures to prevent attribution and that *"[t]ypically such measures **limit** access to the retained additional information (e.g. keys or tables of pseudonyms), and **control** the flow of pseudonymised data."* (emphasis added). If the additional information that might facilitate re-identification is publicly available (as discussed at paragraph 21), it would not be subject to direct *control* by the controller. In such cases, if the controller limits access to the pseudonymised data and prohibits its combination with publicly available data (i.e. takes robust steps to prevent the publicly available data from entering the pseudonymisation domain through, for example, technical, organisational or contractual measures), it is not clear whether this would be deemed to be the exercise of sufficient *control* of the additional information by the controller for the purposes of pseudonymisation.

We recommend that the Guidelines explicitly acknowledge that where the controller is able to limit access to the pseudonymised data and prohibit its combination with publicly available data in this way, the data should be considered pseudonymised. The Guidelines should also provide illustrative examples of such measures, demonstrating how a controller can sufficiently control the risk of re-identification in these circumstances.

Transfer of pseudonymised data to third countries (paragraph 64 and 65)

Where pseudonymisation is used as a supplementary measure to ensure compliance with Article 44 and 46(1) of the GDPR, the *"design of the pseudonymisation procedure needs to start from an assessment of which information the public authorities of the recipient country can be expected to process or to be able to obtain with reasonable means, even if those means may infringe the legal norms in the third country. This information must then be assumed to be available in the pseudonymisation domain."*

Assessing what legal and practical means foreign public authorities could use to obtain access to information can be a challenging and sometimes speculative exercise (e.g. because such laws or practices may not be well developed or documented, or they may not be transparent). It will also be very difficult for a controller to know what information a foreign public authority already has in its possession about a data subject, which it could use to reverse the pseudonymisation (e.g. where a data subject has visited or lived in the third country).

This assessment is made more impractical by the wording of paragraph 65 which requires the reasonable means employed by public authorities to be assessed *"even if those means may infringe the legal norms in the third country"*. This would effectively require the controller to carry out the assessment on the basis that public authorities may contravene the laws they are subject to in order to access information (for which there may be no robust and publicly available evidence).

The Guidelines also do not link the assessment of pseudonymisation (as a supplementary measure for international transfers) with the broader transfer impact assessment process and the fact that a more holistic assessment of transfer risk must be carried out that considers the impact of implementing a range of security and supplementary measures (of which pseudonymisation may form one part).

These factors not only reduce the meaningfulness of the assessment as a whole but they also make completing the assessment potentially challenging for organisations.

We would therefore urge the EDPB to reconsider this criteria and instead focus on other more practical, proportionate and cost effective methods for assessing the effectiveness of the pseudonymisation (e.g. the use of specific high standard pseudonymisation techniques or PETs).

Data subject rights (paragraph 78)

The Guidelines acknowledge that in practice it may not be possible for a controller that receives pseudonymised data to re-identify data subjects from it, and it will not therefore be able to comply with any data subject rights request in respect of that data. However, paragraph 78 then goes on to say that *"...if the data subject can provide the pseudonym or pseudonyms under which data relating to them is stored, and proof that those pseudonyms pertain to them, the controller should be able to identify the data subjects. In consequence, the data subject rights should apply in this case."*

It is not clear how easily in practice a data subject would be able to obtain the pseudonyms that are required to re-identify the data. Presumably, this would involve the data subject first exercising their right of access against the controller that originally pseudonymised the data (to be provided with a copy of the pseudonyms used), and then subsequently providing the pseudonyms to the controller that received the pseudonymised data. This would appear quite impractical (particularly if the original controller relies on exemptions available to it under data protection law to refuse disclosure of the pseudonyms). This avenue may not be an option at all where the pseudonymising controller is not subject to GDPR (e.g. where it is located outside the EEA in a jurisdiction in which there is no right of access under local law).

Requiring the controller that receives pseudonymised data to re-identify it for the purposes of responding to data subject rights requests may also severely undermine the underlying reasons for which that controller was given pseudonymised data in the first place. For example, where a controller has applied effort, care and resources to ensure highly confidential or sensitive information about data subjects (e.g. sensitive health information, or information about criminal convictions) is effectively pseudonymised, so that it can be processed by a second controller (e.g. a services provider), the fact that the second controller could be compelled to re-identify the data (if provided with pseudonyms by the data subject) would undermine the confidentiality of the data.

A further issue with this requirement would be the inconsistency of paragraph 78 of the Guidelines with the AG Opinion. As per the AG Opinion, if the receiving controller is not deemed to be processing personal data at all, then it is not clear why it should be obliged to re-identify the data if the data subject provides it with pseudonyms (i.e. as Article 11(2) of the GDPR would not apply to its processing of pseudonymous data).

We would therefore respectfully suggest that the EDPB should reconsider the requirement in paragraph 78.

Meaning of *identify* within the context of pseudonymisation (paragraph 131)

Controllers are required to consider *"which attributes contained in the personal data that is to be pseudonymised can be used alone or in combination to identify the data subjects directly (identifiers)."* It is not clear what "identify" means within the context of pseudonymisation. For example, in the context of the anonymisation analysis, the risk of a person being "identified" is assessed by considering the risks of "singling out", "linkability" and "inference." However, these concepts do not entirely translate to the pseudonymisation context because pseudonymous data may contain unique pseudonymous identifiers, which clearly allow for "singling out," even if they do not otherwise allow for linkage to other data or inference about the data subject. It would be helpful if the EDPB could provide further clarity in this regard.

Illustrative examples of the application of pseudonymisation (Annex)

The example case studies of pseudonymisation set out in the Annex are specific to very particular kinds of processing (e.g. many of them relate to processing of health / medical data in specific circumstances). These examples may be of limited use to controllers not operating in the industries / scenarios covered. It would be helpful if the EDPB could expand the case studies to a wider range of industries and use cases.