Connect Europe response to the EDPB's Guidelines 01/2025 on Pseudonymisation
March 2025

Connect Europe welcomes the opportunity to respond to the European Data Protection Board's Guidelines 01/2025 on Pseudonymisation and the intention to provide clarification on the use and benefits of pseudonymisation. In this regard, we would like to raise some additional remarks about the guidelines.

General Comments
Connect Europe appreciates the intention of the EDPB Guidelines to clarify the use and benefits of pseudonymisation for data controllers and processors. We also welcome the inclusion of the annex containing several examples highlighting the benefits of pseudonymisation, in light of GDPR-relevant principles.

The guidelines aim at clarifying two key points:

1) Pseudonymised data is personal data, and

2) Pseudonymisation can reduce security risks and facilitate the use of legitimate interest as a legal basis for data processing (Article 6.1.f of the GDPR).

Further clarification on these two points is necessary, especially considering the recent **Opinion of ECJ Advocate General in Case-413/23** which considers that **pseudonymized data can fall outside the concept of "personal data"** for a recipient of the data when it is virtually impossible for the recipient to identify any data subjects from the data (even if it would be possible for the sender of the information).

The ECJ Advocate General considers **whether pseudonymous data may, under certain conditions, fall outside the scope of the concept of "personal data"**, for instance, when pseudonymisation is robustly secure (Par. 51-59 Advocate General's Opinion).

Considering the significance of this case, **the EDPB should wait until the forthcoming ECJ Ruling**, in order to incorporate the conclusions of the Ruling into the Guidelines.

When the notion of pseudonymization was included in the GDPR, the Legislator stressed that the "GDPR would encourage privacy-friendly techniques to reap the benefits of data innovation while protecting privacy". Indeed, the risk reduction resulting from pseudonymisation may enable controllers to rely on legitimate interests under Art. 6(1)(f) GDPR as the legal basis for their processing, provided they meet the other requirements and contribute to establishing compatibility of further processing according to Art. 6(4) GDPR.

However, **so far, telecom companies cannot benefit from Article 6.4. GDPR**, as it is not recognised in the outdated sector specific ePrivacy Directive (Directive 58/2002/EC lastly reviewed by Directive 136/2009/EC). Compatibility through pseudonymization of traffic and location data could help process this data in a secure manner, in order to improve service quality and security or conduct research to advance networks and technologies, all while protecting user privacy. Telecom companies, confronted with an outdated sector specific privacy framework, are prevented from benefitting from the GDPR's Risk-Based Approach.

Pseudonymisation is a key tool for advancing secure AI more quickly.

Data labelling is crucial for pseudonymisation and for auditing. For instance, telecom companies like Telefónica label all direct identifiers and quasi-identifiers of each "data entity", which allows control, audit and automatization, simplifying processes and guaranteeing security.

Certain points remain unclear

There are still several technical and legal points that should be clarified by the guidelines. **Some new concepts are introduced**, which should be further expanded upon and do not have a clear demonstrable advantage. Additionally, there is a **lack of consistency** in the terminology that is used, which does not fully align with concepts used by ENISA[1]. We also find it unfortunate that the text does not mention codes of conduct and certification encouraged by the GDPR (Articles 40-43), which would be helpful for data controllers and processers to demonstrate their responsibility.

A risk-based approach is fundamental when using pseudonymisation, as is taking into account state-of-the-art technology. A case-by-case approach is therefore needed, depending on the context.

<u>New concepts introduced by the Guidelines</u>

Several new concepts are proposed by the guidelines. Some of these concepts could use further clarification, along with further explanation as to how they would be implemented. This is particularly needed in the absence of guidelines on anonymisation or guidelines on the application of security measures according to a risk-based approach in the context of pseudonymisation, where the balance between security and utility must be taken into account.

One of these new concepts is a **"pseudonymisation domain"**. Its goal seems to be to limit the concept of pseudonymisation to an area defined by the recipients of the pseudonymised data. This choice has the advantage of avoiding a context and a risk assessment, however, we are of the opinion that this concept should not be limited to a simple description of the recipients of the pseudonymised data. This could result in decisions being made without a suitable risk analysis method to take into account the risks associated with unauthorized inversion of the pseudonymisation.

The guidelines also introduce other technical terms related to new categories of personal data, such as **"quasi-identifier"**, a concept that is introduced without mentioning where it originated (i.e. OECD, NIST…). The introduction of such terms could have an impact on the classification of personal data and on levels of sensitivity. In general, It would be useful for the EDPB to produce a technical glossary for concepts that are not strictly taken from the GDPR.

<u>Differentiation between pseudonymisation and anonymisation</u>

The document should clarify the difference between pseudonymisation and anonymisation more strongly. The guidelines could mention existing references or future guidelines on anonymisation; for instance, the 2014 WP29 opinion or the anticipated update of the 2014 WP29 opinion.

<u>Connection to other guidelines</u>

[1] https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices

There are several references to technical guidelines, such as those issued by ENISA, but they are sometimes described in a more simplistic manner, sometimes resulting in unclear terminology. To avoid any confusion or misunderstanding, it would be useful to provide explanations of how the text corresponds to other texts that are cited.

Please find below a table that details our specific concerns related to the different sections of the guidelines.

Page (P) Paragraph (N)	Text	Comments	Proposals
General		At no point does the text mention the principles of codes of conduct and certification with regard to Section 5 of the GDPR (in particular Article 40-2-d and Article 42), which would certainly be of interest for data controllers and how they demonstrate their responsibility.	Add a section complying with the principle of article 40-2-d of the GDPR Section 5 Codes of conduct and certification Article 40 Codes of conduct d) the pseudonymisation of personal data; Supporting the principle of certification under Article 42 and measures to encourage its development.
Executive Summary (page 3, 4th paragraph)	Even if all additional information retained by the pseudonymising controller has been erased, the pseudonymised data can be considered anonymous only if the conditions for anonymity are met.	This is useful for highlighting the possibility of obtaining anonymity. For example, if the key is deleted, it may be possible to obtain anonymisation in certain cases. Anonymity is often confused with pseudonymity. It would be useful to reiterate what anonymity is, not to be confused with pseudonymity.	Reiterate the conditions for anonymisation. Refer to the 2014 WP29 opinion or the replacement EDPS text currently being drafted. Pseudonymisation is not anonymisation. Clarify the difference between pseudonymisation and anonymisation

		Same point for page 10 N22	
Executive Summary (page 3, 7th paragraph)	Finally, the contribution of pseudonymisation to data protection by design and default, and the assurance of a level of security appropriate to risk may make other measures redundant – even though pseudonymisation alone will normally not be a sufficient measure for either.	Unclear, phrased in a negative way.	Find a more positive way to emphasize the conditions for implementing privacy by design principles, to make complementary security measures operational and adapted to the risk analysis, depending on the level of protection required and the usefulness of pseudonymised data.
Executive Summary (page 3, 8th paragraph)	Define the risks	Are these the risks defined in 2.2.1? Shouldn't we be more explicit about the notion of "risks" and the methodology that would be used?	The risk-based approach is fundamental. A reminder of this risk-based approach would be desirable, referring to articles of the GDPR (recitals 71, 83, articles 24, 25, 32) and ENISA documents, for example. A risk analysis methodology adapted to the implementation of pseudonymisation should take into account the risks associated with the unauthorized reversal of pseudonymisation, depending on the security techniques chosen and data utility requirements.
Executive Summary (Page 4, 1st paragraph)	This context will be called the *pseudonymisation domain* in these	Introducing a new concept*: pseudonymisation domain*	Evaluate the impact of this concept on the risk-based approach and on actor mapping.

	guidelines.		Here, the security perimeter is defined in terms of protection against the actors who will process the pseudonymised data. The definition should take into account all threats within a limited security perimeter, and adapt measures with a required level of protection and a response adapted to the usefulness of pseudonymised data, and not just a mapping of recipients of pseudonymised data.
Page 8 N10	The guidelines introduce a new concept, called pseudonymisation domain, to capture one aspect of that freedom: to determine who should be precluded from attributing the pseudonymised data to individuals.	What kind of "freedom" are we referring to here?	"freedom" to be explained here
Page 8 N11	The Guidelines highlight the benefits of pseudonymisation.	It's useful to have a document that highlights pseudonymisation!	Long-awaited document showing the benefits and advantages of this measure in GDPR-compliant conditions.
P9 N18	Pseudonymising controllers	New concept Avoid confusion with data controller Pseudonymisation may be carried out by a subcontractor or a data subject. The use of the term "controller" may	Check that this concept does not cause confusion with the term "controller" defined by the GDPR.

		lead to confusion with the concept of "data controller". ENISA, for example, uses the term "pseudonymisation entity" to describe a "pseudonymisation entity".	
Page 10 N22	Pseudonymised data, which could be attributed to a natural person by the use of additional information, is to be considered information on an identifiable natural person,7 and is therefore personal. This statement also holds true if pseudonymised data and additional information are not in the hands of the same person. If pseudonymised data and additional information could be combined having regard to the means reasonably likely to be used by the controller or by another person, then the pseudonymised data is personal. Even if all additional information retained by the pseudonymising controller has been erased, the	The document should more clearly differentiate between pseudonymisation and anonymisation. It could mention existing references or future guidelines on anonymisation; for instance, the 2014 WP29 opinion or the text replacing the 2014 WP29 opinion.	Note the reference to Recital 26 of the GDPR on the notion of anonymous information, i.e. information not relating to an identified or identifiable natural person and personal data rendered anonymous in such a way that the data subject is not or is no longer identifiable. This Regulation does not, therefore, apply to the processing of such anonymous information, including for statistical or research purposes.

	pseudonymised data becomes anonymous only if the conditions for anonymity are met.		
Page 10 N26	In accordance with Rec. 28 GDPR, pseudonymising data reduces risks for data subjects while allowing for general analysis.	Recital 26 mentions risks for data subjects, but also assistance to controllers and processors. Why limit the reference to data subjects without mentioning assistance for controllers and processors? "The application of pseudonymisation to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations."	"help controllers and processors to meet their data-protection obligations" should be mentioned in Recital 28. Paragraph 30 clearly states "and the benefits the controllers may derive from it".
Page 11 N31	usefully analysed	Pseudonymisation is a process that ensures data security while fully preserving its usefulness.	Introduce the notion of "utility", often used in reference texts [ENISA19] for pseudonymisation and the compromise between protection and utility.
Page 12 N35 & N36	Pseudonymisation domain	The definition of this context-related concept should take into account all threats, and not just the mapping of pseudonymised data recipients.	Avoid confusing the context of risk analysis with a simple mapping of pseudonymised data recipients.
Page 12 N37	Pseudonymising controller	The entity responsible for pseudonymisation may be a data controller, a processor (performing	Avoid confusion over the term "controller" and prefer the term "entity" (as ENISA uses).

		pseudonymisation on behalf of a data controller), a trusted third party or a data subject, depending on the pseudonymisation scenario.	
Page 13 N44	Building block	An example in ANNEX to illustrate this concept would be useful.	Add an example of a set of measures to be implemented in the appendices
Page 13 N45	Protection by design	In accordance with Article 25(3), the certification of pseudonymisation mechanisms under Article 42 may be used to demonstrate compliance with data protection by design and data protection by default.	Mention certification with reference to Article 25 Guidelines 4/2019 Data protection by design and data protection by default Version 2.0 Adopted October 20, 2020
Page 14 N47	Pseudonymised "consistently"	Concept of "consistently" to be clarified	Add the more commonly used "Deterministic pseudonymisation" concept
Page 14 N49	Data protection by default	Objective already mentioned in N45	Remove redundant references
Page 14 N52	"Group of collaborating controllers" and "participating controllers".	Is this the same concept as "joint controllers" as defined in article 26?	Replace with "joint controllers" or justify the choice of a different term
Page 15 N56	"compatible purposes"	Reference to recital 50	Refer to recital 50
Page 16 N63		No mention of certification	Possible reference to Guidelines 07/2022 on certification as a tool for transfers. The GDPR places considerable trust in private certification mechanisms as a "regulated self-regulation".
Page 19 N79	Indicate in the information	This principle may be difficult to implement	Specify when and how to inform (information

		and may run counter to the principle of security (for example, in relation to the confidentiality of pseudonymised data processing within a company, in order to separate types of processing according to the data access rights granted to employees).	leaflet)
Page 21 N87	"Lookup tables" or "mapping tables" or "tables matching" or "table linking" depending on the passages in the text.	The notion of correspondence tables is noted in different ways. • "Mapping tables" by ENISA. • lookup tables" or "tables matching" in the text.	Standardizing the use of this concept Replace with "mapping tables" referring to ENISA technical documents Alternatively, propose a uniform term and specify this notion in a "technical" glossary.
Page 23-24	"Quasi-identifiers" “Pseudonymisation proxy” "randomly generated pseudonyms”	For all technical vocabulary, a glossary and comparison with references would be useful.	Propose a technical glossary for concepts not mentioned in the GDPR and mention references
Page 23 N101	"Quasi-identifiers"	The new concept of personal data, for which no reference is mentioned (OECD, NIST?), could have an impact on the classification of personal data and qualifying the sensitivity levels of this data. Is this one of the aims of the text?	Specify reference. Measure the impact of this concept in relation to the GDPR's definition of personal data.
Page 25 N109	"vetted"?	What does the notion of "vetted" correspond to? Could other	Other concepts to add: Auditable, certified, certificate, codes of

		objectives set by the GDPR such as codes of conduct or certification be mentioned at this level?	conduct ...?
Page 26 N116 & N117	"person pseudonyms" "relationship pseudonyms"	Some concepts are taken from scientific articles, in particular article 33. Do they have the same meaning?	Clarify terms already used in the academic world to maintain consistency
Page 27 N119	"interaction of a vehicle with an intelligent transport".	Example of a very specific "interaction of a vehicle with an intelligent transport"	Place specific examples in appendix
Page 30 N131	"to assess the risk of attribution"	A definition of the notion of "risk of attribution" would seem useful. Re-defining the concept of an "attack"? "Recovery"?	Specify this point as important. Refer to methodology?
Page 31 Example 1	User / patient / data subject	Specify user / patient / data subject	Specify that the user is a patient, i.e. the person concerned
Page 45 Glossary	"Quasi-identify" "Lookup table" "Pseudonymisation" "proxy" "person " "pseudonym"...	New concepts are introduced in the guidelines with references. If terms are defined specifically for the guidelines, clarification would be useful.	Enrich the glossary with all the terms that are complementary to the GDPR and essential for understanding and implementing pseudonymisation. Cite references if the terms have known sources.

Comparative table of terms mentioned

terms	EDPB definition	Other references
pseudonymisation domain	Environment in which the controller or processor wishes to preclude attribution of data to specific data subjects. May incorporate persons acting under the authority of the controller or processor, respectively, other natural or	?

	legal persons, public authorities, agencies or other bodies, and their respective technological and informational resources. Does not include persons authorised to process additional data allowing the attribution of the pseudonymised data to data subjects.	
Pseudonymising controller or processor	These guidelines will call controllers that use pseudonymisation as a safeguard and modify original data according to Art. 4(5) GDPR pseudonymising controllers. Similar terminology is used for processors.	[ENISA19] Pseudonymisation entity is the entity responsible of processing identifiers into pseudonyms using the pseudonymisation function. It can be a data controller, a data processor (performing pseudonymisation on behalf of a controller), a trusted third party or a data subject, depending on the pseudonymisation scenario. It should be stressed that, following this definition, the role of the pseudonymisation entity is strictly relevant to the practical implementation of pseudonymisation under a specific scenario. However, in the context of this report, the responsibility for the whole pseudonymisation process (and for the whole data processing operation in general) always rests with the controller.
Lookup tables", table matching or table linking	Such additional information may consist of tables matching pseudonyms with the identifying attributes they replace. procedures that create lookup tables matching identifiers with the pseudonyms used to replace them.	Pseudonymisation mapping table is a representation of the action of the pseudonymisation function. It associates each identifier to its corresponding pseudonym. Depending on the pseudonymisation function P , the pseudonymisation mapping table may be the pseudonymisation secret or part of it. [ENISA19]

Quasi-identifier	One way to attribute data to a natural person is by looking at several attributes contained in the data that reveal information about the physical, physiological, genetic, mental, economic, cultural or social identity of the data subject. If a combination of those attributes are sufficient to attribute at least part of the pseudonymised data to data subjects, then they are called quasi-identifiers.	Specify if this concept refers to other guides? Which reference?
Pseudonymisation proxy	All relevant incoming data is first processed by a dedicated, separate team. The persons authorised to reverse pseudonymisation (Rec. 29 GDPR, second pseudonymisation is reversed, and the original collected data turned over for processing.	?
Pseudonymisation at the source	Pseudonymisation is already performed by the controller that is the source of the information, prior to transmission to the entity processing the pseudonymised data.	?
randomly generated pseudonyms	When using lookup tables for the pseudonymising transformation, it suffices to choose randomly generated pseudonyms. When using cryptographic algorithms, suitable building blocks include (keyed) pre-image resistant29 cryptographic one-way functions (like HMACs) or encryption schemes guaranteeing cipher text indistinguishability30 (like symmetric block ciphers used in a suitable mode).	[ANON] Identifiers which are generated using random data only, i.e., fully independent of the subject and related attributes, do not contain side information on the identified subject, whereas non-random identifiers may do. E.g., nicknames chosen by a user may contain information on heroes he admires; a sequence number may contain information on the time the pseudonym was issued; an e-mail address or phone number contains information how to

		reach the user.
person	One or several controllers may choose to pseudonymise all data they process relating to the same data subjects consistently. The corresponding pseudonyms are usually called person pseudonyms.	[ANON] A person pseudonym is a substitute for the holder's name which is regarded as representation for the holder's civil identity. It may be used in all contexts, e.g., a number of an identity card, the social security number, DNA, a nickname, the pseudonym of an actor, or a cell phone number.
role		[ANON] role pseudonym: The use of role pseudonyms is limited to specific roles51, e.g., a customer pseudonym or an Internet account used for many instantiations of the same role "Internet user". The same role pseudonym may be used with different communication partners. Roles might be assigned by other parties, e.g., a company, but they might be chosen by the subject himself/herself as well.
relationship	A controller may also choose to pseudonymise all data consistently that it intends to process for one or several particular purposes defining a certain type of relationship of data subjects with that controller. For instance, a data subject may be assigned different pseudonyms depending on whether the data concern their relationship with controllers as employees or customers. In this case, pseudonymisation secrets (or parts thereof) are maintained	[ANON] For each communication partner, a different relationship pseudonym is used. The same relationship pseudonym may be used in different roles for communicating with the same partner. Examples are distinct nicknames for each communication partner

	only for the time the relationship with the data subject lasts. The resulting pseudonyms are called relationship pseudonyms. The use of such pseudonymisation is only admissible if linking of different pieces of pseudonymised data relating to the same person in the same relationship to the controller may become necessary and will be lawful in this case. This condition is often fulfilled if there is only one common purpose, or the various purposes are compatible.	
role-relationship pseudonym		[ANON] role-relationship pseudonym: For each role and for each communication partner, a different role-relationship pseudonym is used. This means that the communication partner does not necessarily know, whether two pseudonyms used in different roles belong to the same holder. On the other hand, two different communication partners who interact with a user in the same role, do not know from the pseudonym alone whether it is the same user