Comments of Latvian Information and Communications Technologies Association (LIKTA) member organizations – ICT companies

First comment

In Latvian language:

Pseidonimizācijas regulējums jau šobrīd ir aktuāls, un paredzams, ka tā nozīme turpinās pieaugt. Pastāv iespēja, ka tas kļūs par galveno metodi drošai mākslīgā intelekta modeļu apmācībai situācijās, kad nepieciešams izmantot datu kopas, kas satur jūtīgus personas datus.

Komentāri par Eiropas Datu aizsardzības koleģijas (European Data Protection Board) pseidonimizācijas vadlīnijām:

1. Vadlīnijas nosaka iespējas un ierobežojumus jutīgu datu apstrādātājiem (procesors) un pārvaldniekiem (controllers) ar mērķi mazināt jutīgu t.sk. personas datu ļaunizmantošanas riskus.

2. No datu apstrādātāju un pārvaldnieku viedokļa ir svarīgi ievērot samērības principus, lai no vienas puses mazinātu datu ļaunizmantošanas riskus, bet no otras puses būtiski neierobežotu tehnoloģiju pētnieku un izstrādātāju iespējas veidot risinājumus, kas pievieno vērtību to izmantotājiem t.sk. arī personām, kuru dati tiek pseidonimizēti šo risinājumu vai to komponentu izstrādes procesā.

3. Šādas līdzsvarotas pieejas nepieciešamība attiecas arī uz Vispārīgās datu aizsardzības regulas prasību piemērošanu un interpretēšanu attiecībā uz dažādām datu aizsardzības metodēm t.sk. pseidonimizāciju.

4. Atbildību par atbilstošas pieejas (t.sk, datu pirmapstrādes metožu) izvēli jāuzņemas datu pārvaldniekam kopīgi ar datu apstrādātājiem, nodrošinot iespējami zemus šādu datu ļaunizmantošanas riskus.

5. Jutīgu t.sk. personas datu izmantošana dažādu tehnoloģisku risinājumu – valodas tehnoloģiju (t.sk. lielo valodas modeļu), datu analīzes un prognozēšanas risinājumu, datorredzes tehnoloģiju u.c. pētniecībā un izstrādē ir kritiska, lai veidotu kvalitatīvus risinājumus, kas nodrošina datos balstītu lēmumu pieņemšanu, analītiku un prognozēšanu, kā arī personalizētu produktu un pakalpojumu izstrādi. Bieži vien tieši datu pseidonimizācija ir vienīgā metode, kas nodrošina šādu uzdevumu veikšanu.

6. Jebkādi ierobežojumi, kā arī pārmērīgas prasības attiecībā uz datu pārvaldnieku un apstrādātāju veicamajiem papildu uzdevumiem jutīgu datu apstrādes procesā, rada papildu izmaksas, kas var kavēt jaunu tehnoloģiju un risinājumu attīstību, kā arī mazināt Eiropas uzņēmumu konkurētspēju, salīdzinājumā ar tiem, kas darbojas apstākļos ar mazākiem ierobežojumiem.

Automated translation:

(The regulation of pseudonymization is already relevant today, and its importance is expected to continue to grow. It is possible that it will become the main method for safe training of artificial intelligence models in situations where it is necessary to use data sets containing sensitive personal data.

Comments about the European Data Protection Board's guidelines on pseudonymization:

1. The guidelines set out the possibilities and restrictions for processors and controllers of sensitive data with the aim of mitigating the risks of misuse of sensitive, including personal, data.

2. From the perspective of data processors and controllers, it is important to observe the principles of proportionality in order to, on the one hand, mitigate the risks of misuse of data, but on the other hand, not significantly limit the possibilities of technology researchers and developers to create solutions that add value to their users, including persons whose data is pseudonymized in the process of developing these solutions or their components.

3. The need for such a balanced approach also applies to the application and interpretation of the requirements of the General Data Protection Regulation in relation to various data protection methods, including pseudonymisation.

4. Responsibility for the selection of an appropriate approach (including data pre-processing methods) must be assumed by the data controller together with data processors, ensuring the lowest possible risks of misuse of such data.

5. The use of sensitive, including personal data, in the research and development of various technological solutions - language technologies (including large language models), data analysis and forecasting solutions, computer vision technologies, etc. is critical for creating high-quality solutions that ensure data-driven decision-making, analytics and forecasting, as well as the development of personalized products and services. Often, data pseudonymisation is the only method that ensures the performance of such tasks.

6. Any restrictions, as well as excessive requirements regarding additional tasks to be performed by data controllers and processors in the process of processing sensitive data, create additional costs that may hinder the development of new technologies and solutions, as well as reduce the competitiveness of European companies compared to those operating in less restrictive conditions.

Second coment

The guidelines adopt an overly restrictive view of pseudonymisation that confuses pseudonymisation and anonymisation, and ignores both the text of GDPR and established CJEU case law.

- At multiple points, the guidelines suggest that, for pseudonymisation to be effective:
    - The “pseudonymisation domain” may in some instances have to be defined as including **any and all third parties** that may theoretically attempt to access the pseudonymised data and additional information, even if they are not authorised to do so; and
        - See, e.g., 21-22, 37-38, 42-43.

    - It must not be possible for any party in the pseudonymisation domain to identify an individual in the pseudonymised data, taking into account **all means reasonably likely to be used**, including accessing information beyond that actually held by the pseudonymising controller and parties with whom the pseudonymised data is shared.
        - In other words, no parties in the pseudonymisation domain should be able to obtain with reasonable efforts any additional information enabling attribution of the pseudonymised data to specific data subjects.
        - See, e.g., 21-22, 42-43, 60.
- But this reasoning suffers from a fatal flaw - it **adopts an overly restrictive review of pseudonymisation, confusing the concepts of pseudonymisation and anonymisation**.
    - If it is not possible for a party to attribute data to an identifiable individual considering all means reasonably likely to be used, then the data is anonymous, not pseudonymous, with regard to that party.
    - Effectively pseudonymising data must be understood as processing data in such a way that strips the data of some information, without which it is not possible to attribute the data to a specific data subject, and which is kept separate and subject to technical and organisational measures.
    - In other words, pseudonymising data **does not require**: (1) considering any and all third parties that may theoretically attempt to access the pseudonymised data and attribute it to individuals; or (2) considering any and all means reasonably likely to be used by parties in the pseudonymisation domain - including additional information that may be accessed - to attribute the pseudonymised data to individuals.
- This is clear from **the text of GDPR**
    - GDPR Art. 4 and Recital 29 make clear that **the pseudonymisation domain** will not have to include any and all third parties that may theoretically attempt to access the data.
        - **GDPR Art. 4** defines “pseudonymisation” as “the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.” This definition implies two important things: (1) the “additional information” at issue is additional information - such as pseudonyms - generated by the pseudonymising controller from an act of processing personal data; and (2) it is exactly “such” additional information - not any other additional information - that must render an individual identifiable if combined with pseudonymised data. This, in turn, means that the pseudonymisation domain cannot extend beyond the controller and parties with whom they have shared pseudonymised data; any such additional parties would not be

        able to attribute the pseudonymised data to specific individuals using the additional information at issue.

        - **Recital 29** also underscores that the pseudonymisation domain should not be understood as potentially including any and all third parties that may attempt to access the data. In particular, Recital 29 states that, for the purpose of incentivising pseudonymisation, pseudonymisation "should be possible within the same controller." If the pseudonymisation domain often required consideration of any and all third parties that may attempt to access the data, then pseudonymisation often would not be possible within the same controller.

    - GDPR Recital 26 makes clear that pseudonymisation cannot require a consideration of all means reasonably likely to be used to re-identify an individual. That test applies to **anonymisation, not pseudonymisation**.

        - **Recital 26** states that if it is not possible "to identify the natural person directly or indirectly" when considering "all the means reasonably likely to be used, either by the controller or by another person," then an individual is not identifiable; the data is anonymous. Data can still be pseudonymous data even if it would be possible to identify an individual considering the means reasonably likely to be used.

- **CJEU case law** also makes clear that assessing whether an individual could be identified directly or indirectly, considering all means reasonably likely to be used, is a legal test for anonymisation, not pseudonymisation.

    - The CJEU in ***Breyer*, *Scania*, and *IAB Europe*** applied that test to determine whether data was anonymous data - not to determine whether data was pseudonymous.

<u>The guidelines should not address anonymisation or the concept of personal data, but they do, and they imply an overly restrictive view of anonymisation that conflicts with EU case law.</u>

- These guidelines are intended to address the concept of pseudonymisation, which has a straightforward, ordinary meaning under GDPR Art. 4.

    - But the guidelines inappropriately extend beyond pseudonymisation, addressing the concepts of anonymisation and personal data both directly and indirectly, partly by confusing pseudonymisation and anonymisation, as described above.

- Not only do the guidelines seem to confuse the concepts of pseudonymisation and anonymisation, but they also seem to advocate for an **overly broad interpretation of personal data and an overly restrictive view of when data are effectively anonymised**.

    - In particular, the guidelines seem to suggest that, when a pseudonymising controller shares pseudonymised data with an authorised third party, that data may not be pseudonymous with respect to the authorised third party if other, unauthorised third parties may attempt to gain access to the data and re-identify individuals using means available to them, but not to the authorised third party. See, e.g., 22, 43.

And, if the data are not pseudonymous, then it follows that the data cannot be anonymous.

  - But in such an instance - where the authorised third party does not have reasonably available means to re-identify individuals - then the data should be properly understood as anonymous, not pseudonymous, with regard to that party.

- This is clear from **case law** of both the CJEU and the EU General Court.
  - In *Breyer*, the **CJEU** emphasized that if "the risk of identification appears in reality to be insignificant" because identification would "require[] a disproportionate effort in terms of time, cost and man-power," then the data is anonymous from the perspective of the party for which identification would be nearly impossible.
  - The **EU General Court** built on the CJEU's *Breyer* ruling in *SRB* to emphasize that the risk of identification must be assessed from the perspective of the party holding the data. The question is not whether any third parties may theoretically be able to identify an individual; it is whether the third party in possession of the data has means reasonably likely to be used by them to identify an individual without disproportionate effort.
    - The CJEU's judgment in the appeal of *SRB* is expected soon, and it may provide binding authority on the issue of anonymisation, which the EDPB should not attempt to preempt in guidelines on pseudonymisation.

<u>The guidelines misunderstand how pseudonymisation interacts with GDPR Art. 11</u>

- We appreciate that the EDPB recognises that the data subject rights of GDPR Arts. 15-20 generally do not apply to pseudonymised data.
- But the guidelines contain **two misunderstandings** about the obligations that GDPR Art. 11 imposes on controllers.
- First, the guidelines misunderstand **when controllers must inform data subjects** about the applicability of Art. 11(1).
  - In particular, the guidelines imply that a controller is subject to this obligation whenever "it holds" pseudonymised data. See 77-79.
  - But this conflicts with the plain text of Art. 11(1), which applies "[i]f the purposes for which a controller processes personal data do not or do no longer require the identification of a data subject by the controller."
  - This text makes clear that a controller has to be engaged in an act of "process[ing] personal data" in the first instance for the obligations under GDPR Art. 11 to apply.

    - If a controller never processes personal data in a particular context—if, in that context, it only ever holds data not requiring identification of a data subject - Art. 11's obligations of informing data subjects do not apply.

- Second, the guidelines misunderstand **what information controllers must provide** to data subjects under Art. 11(2) (if and when they are obligated to).
    - In particular, the guidelines state that controllers should inform data subjects "how they can obtain the pseudonyms relating to them, and how they can be used to demonstrate their identity. In this case, the controller may need to provide the identity and the contact details of the source of the pseudonymised data or of the pseudonymising controller." See 79.
    - But this goes far beyond what the text of Art. 11(2) requires. Art. 11(2) states only that, "[w]here . . . the controller is able to demonstrate that it is not in a position to identify the data subject, the controller shall inform the data subject accordingly, if possible."
        - In other words, if it is "possible"- not always - the controller should inform the data subject merely that it cannot identify the data subject - nothing more.
    - Further, the guidelines' suggestion that controllers should provide data subjects with pseudonyms directly conflicts with the text of Art. 11(2), which states that it is the responsibility of the data subject to "provide additional information enabling his or her identification."

Taken together, the guidelines' infirmities will have the effect of disincentivising privacy-preserving practices like pseudonymisation and anonymisation.

- The guidelines rightfully recognize that privacy-preserving practices like pseudonymisation are valuable and should be incentivized.
    - Such practices help not only to enhance individuals' privacy, but also to enable data sharing practices that drive technological advancement and economic growth.
- But the guidelines unfortunately suggest that it will be very difficult in practice to prove that data are pseudonymous, let alone anonymous.
    - Not only do the guidelines adopt overly restrictive views of pseudonymisation and anonymisation, but they also contain unnecessarily complicated technical discussions suggesting that achieving pseudonymisation will be technically challenging in practice.
- Being able to show that data are pseudonymous or anonymous is a powerful incentive for organisations to innovate and invest - one that these guidelines severely limit.