1-Digital Omnibus on GDPR:

The EU Commission presented a package of proposed amendments which includes changes to the EU General Data Protection Regulation (GDPR), the Data Act and other digital laws (Digital Omnibus).

1. Definition of personal data (GDPR Article 4)

Current GDPR:

Article 4: personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Recital 26: The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person.

Amendments to Regulation (EU) 2016/679 (GDPR)

Regulation (EU) 2016/679 is amended as follows:

1.Article 4 is amended as follows:

(a)in point 1, the following sentences are added:

'Information relating to a natural person is not necessarily personal data for every other person or entity, merely because another entity can identify that natural person. Information shall not be personal for a given entity where that entity cannot identify the natural person to whom the information relates, taking into account the means reasonably likely to be used by that entity. Such information does not become personal for that entity merely because a potential subsequent recipient has means reasonably likely to be used to identify the natural person to whom the information relates.'

Commission's explanation regarding the amendment: Paragraph 1 would clarify the definition of personal data under Article 4 of Regulation (EU) 2016/679 (General Data Protection Regulation) by stating that information is not to be considered personal data for a given entity when it does not have means reasonably likely to be used to identify the

natural person to whom the information relates. As a result, such an entity would not, in principle, fall within the scope of application of that Regulation.

What does that mean?

The proposed new package aims to replace this "objective" definition (which others can also define) with a "subjective" (controller-specific) definition that only considers the capabilities and intentions of the data processing organization.

The risk is that, the level of protection would vary from "company to company" based on their own claims of identification capability. This could allow companies to process sensitive information without safeguards, simply by claiming they don't know who the person is.

GDPR Article 4(5): pseudonymisation' means 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.

Normally, as mentioned above, current GDPR Recital 26 says that 'The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an **identifiable natural person.'**

However, Proposed GDPR Recital 27 says that 'The existence of additional information enabling the data subject to be identified does not, in itself, mean that pseudonymised data must be regarded as constituting, in all cases and for every person or entity, personal data for the purposes of the Article 4(1) | Definition of "Personal Data" 6 which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.'

In the GDPR framework, for a piece of information to be protected as **"Special Category/Sensitive Data (health data)" (Article 9)**, it must first meet the definition of **"Personal Data" (Article 4(1))**. Article 4(1) is the "gateway" to the entire regulation.

Current Situation (Objective Approach): If *anyone* (the hospital that pseudonymized the data) can identify the person, the data remains "pseudonymized personal data." Therefore, even for a third-party company (Entity B), it is still considered "health data" and is subject to the strict protections of Article 9.

Proposed Change (Subjective Approach): The definition becomes "entity-specific." If Entity B claims, "With the reasonable means available to me, it is impossible to identify

this person," then that data is **no longer personal data** specifically for Entity B. If it's not personal data, it cannot be "special category/health data" either.

Conclusion: Normally, processing health data is generally prohibited unless specific exceptions apply. However, with the proposed article, in this case since the data is no longer "personal," the company can analyze it, sell it, or use it to train AI without following GDPR's strict sensitivity rules.

2- Special-category data (GDPR Article 9)

Current GDPR: Processing of special categories of personal data – Article 9

1-Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, **data concerning health** or data concerning a natural person's sex life or sexual orientation shall be prohibited.

2-Paragraph 1 shall not apply if one of the following applies: (exceptions)

.....

Amendments to Regulation (EU) 2016/679 (GDPR)

Proposed text:

in paragraph 2, the following points are added: (to the exception part)

'(k) processing in the context of the development and operation of an AI system as defined in Article 3, point (1), of Regulation (EU) 2024/1689 or an AI model, subject to the conditions referred to in paragraph 5.

(b)the following paragraph is added:

'5. For processing referred to in point (k) of paragraph 2, appropriate organisational and technical measures shall be implemented to avoid v the collection and otherwise processing of special categories of personal data. Where, despite the implementation of such measures, the controller identifies special categories of personal data in the datasets used for training, testing or validation or in the AI system or AI model, the controller shall remove such data. If removal of those data requires disproportionate

effort, the controller shall in any event effectively protect without undue delay such data from being used to produce outputs, from being disclosed or otherwise made available to third parties.'

Comission's explanation – Recital 33: The development of certain AI systems and AI models may involve the collection of large amounts of data, including personal data and special categories thereof. Special categories of personal data may residually exist in the training, testing or validation data sets or be retained in the AI system or the AI model, although the special categories of personal data are not necessary for the purpose of the processing. In order not to disproportionately hinder the development and operation of AI and taking into account the capabilities of the controller to identify and remove special categories of personal data, derogating from the prohibition on processing special categories of personal data under Article 9(2) of Regulation (EU) 2016/679 should be allowed. The derogation should only apply where the controller has implemented appropriate technical and organisational measures in an effective manner to avoid the processing of those data, takes the appropriate measures during the entire lifecycle of an AI system or AI model and, once it identifies such data, effectively remove them. If removal would require disproportionate effort, notably where the removal of special categories of data memorised in the AI system or AI model would require re-engineering the AI system or AI model, the controller should effectively protect such data from being used to infer outputs, being disclosed or otherwise made available to third parties. This derogation should not apply where the processing of special categories of personal data is necessary for the purpose of the processing. In this case, the controller should rely on the derogations pursuant to Article 9(2)(a) – (j) of Regulation (EU) 2016/679.

What does that mean?

Basically, Recital 33 states that, since AI models are trained on massive datasets, **sensitive data like health records** or political opinions can get mixed in by mistake or unintentionally. This article states:

1. Normally, processing this data is forbidden, but we are leaving an open door (exception) so that AI development isn't blocked.
2. If a company sees this data, they must delete it.
3. However, if they argue, "We would have to tear down and rebuild the entire system, deletion is impossible" (disproportionate effort), then instead of deleting it, they can settle for preventing that data from affecting the AI's results or leaking to others.

That means, The Commission indicates that accidentally obtaining sensitive data (health, religion, etc) during AI acquisition training and implementation constitutes an exemption under Articles 9(2)(k) and 9(5).

The text allows a company to keep data if deleting it requires disproportionate effort (if the AI model would need to be retrained from scratch). However, the concept of "disproportionate effort" is extremely elastic. Companies can always claim that retraining a multi million dollar model is an unreasonable burden. This leads to sensitive data being permanently trapped within the models.

The text suggests: "Do not delete the data, but do not use it to produce output." The problem is that, AI models do not store data like a traditional database, they store information as weights and parameters. How can we know whether a specific piece of health data was "not used" to generate a specific output? Companies cannot fully guarantee this technically.

Normally, the GDPR requires that any data processing must be necessary for a specific purpose. The problem is that, Recital 33 explicitly allows data to remain in a system "even though it is not necessary" for the purpose of processing. This directly contradicts the core GDPR pillars of data minimization and necessity.

In conclusion, only data that **directly reveals health,** political opinions, religion, or sexuality would count as special category data. Inferred information would no longer receive extra protection. Inferred information would no longer receive extra protection. Currently, if an algorithm analyzes your behavior and predicts that you are pregnant, ill, or hold specific political beliefs, that prediction is treated as Sensitive Personal Data (special category data) and is strictly protected. The new proposal seeks to remove this extra layer of protection for traits that are "inferred" rather than directly provided.

Algorithms use behavior patterns to make sensitive predictions about a person's psychological state or health.

- The Problem: Companies could use these "guesses" to change prices (dynamic pricing) or filter candidates out of a hiring process.
- The Consequence: Because this wouldn't be legally classified as "processing health data" under the new rules, it becomes nearly impossible for regulators to monitor or prevent discrimination based on these hidden traits.

Example of Lindenapotheke case: the Court of Justice of the EU ruled that even ordering non-prescription pharmacy medicines counts as health data if it allows conclusions to be drawn about a person's health.

The Omnibus proposal would weaken this reasoning and legitimize processing practices that expose people's vulnerabilities.

3- Legitimate Interest Article 88c:

Proposed article 88c:

Where the processing of personal data is necessary for the interests of the controller in the context of the development and operation of an AI system as defined in Article 3, point (1), of Regulation (EU) 2024/1689 or an AI model, such processing may be pursued for legitimate interests within the meaning of Article 6(1)(f) of Regulation (EU) 2016/679, where appropriate, except where other Union or national laws explicitly require consent, and where such interests are overridden by the interests, or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

Any such processing shall be subject to appropriate organisational, technical measures and safeguards for the rights and freedoms of the data subject, such as to ensure respect of data minimisation during the stage of selection of sources and the training and testing of AI an system or AI model, to protect against non-disclosure of residually retained data in the AI system or AI model to ensure enhanced transparency to data subjects and providing data subjects with an unconditional right to object to the processing of their personal data.

What does it mean?

Article 88c aims to bring the development and operation of AI systems directly under the scope of legitimate interest (GDPR Article 6(1)(f)).

While Article 88c constructs a "legitimate interest" highway for general personal data (Article 6), it leaves health data vulnerable in two primary ways:

1. The Bridge Built by Article 9(2)(k) and Inferred Data: The proposed Article 9(2)(k) lifts the prohibition on processing special categories of data within the context of AI development, particularly where such data is "residual" or necessary for system operation. When combined with Article 88c, this creates a significant loophole for inferred health data.

Instead of processing direct medical records, companies can process large-scale datasets under "legitimate interest" to train AI. If the AI then "infers" a patient's health condition (for example detecting a systemic disease from a dental scan), the company may argue that this prediction is a mathematical output of the system's operation rather than a direct processing of health data.

By reclassifying these sensitive insights as "AI inferences" rather than "special category data," companies effectively use the bridge between Article 9(2)(k) and 88c to bypass the strict consent requirements of the current GDPR. This shifts the focus from what the data is to how the AI uses it, leaving the most sensitive predictions unregulated.

2. Data Retention via "Disproportionate Effort" When Article 9(5) and Article 88c are read together, they create a permanent loophole. If health data is "accidentally" mixed into an AI model and removing it requires a "disproportionate effort" (such as retraining the entire model from scratch), that data is allowed to remain within the system. This means sensitive, irreversible health data could become permanently trapped inside commercial models under the excuse that it is "too difficult to delete." This effectively bypasses the core GDPR principles of "purpose limitation" and "data minimization".

2-Digital Omnibus on AI Act:

Current AI Act: Article 10: Data and Data Governance

5-To the extent that it is strictly necessary for the purpose of ensuring bias detection and correction in relation to the high-risk AI systems in accordance with paragraph (2), points (f) and (g) of this Article, the providers of such systems may exceptionally process special categories of personal data, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons. In addition to the provisions set out in Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, all the following conditions

Amendment to the Regulation 2024/1689

the following Article 4a is inserted in Chapter I:

Article 4a

Processing of special categories of personal data for bias detection and mitigation

1. To the extent necessary to ensure bias detection and correction in relation to high-risk AI systems in accordance with Article 10 (2), points (f) and (g), of this Regulation, providers of such systems may exceptionally process special categories of personal data, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons. In addition to the safeguards set out in Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, as applicable, all the following conditions shall be met in order for such processing to occur:

What does it mean?

The General Rule: Prohibition of Sensitive Data Under Article 9(1) of the GDPR, the processing of special categories of personal data (such as ethnic origin, political opinions, religious beliefs, health data, and biometric data) is, as a rule, prohibited. To bypass this prohibition, a controller must typically obtain the individual's explicit consent or rely on one of the specific derogations listed in Article 9(2) (public health, legal obligations, or vital interests).

The Current Framework: The AI Act's Exception Currently, Article 10(5) of the AI Act provides a specific exception for high-risk AI systems. It allows the processing of sensitive data for the purpose of "bias detection and correction," even without consent, by relying on Article 9(2)(g) of the GDPR ("substantial public interest"). However, this is subject to a very high threshold, the processing must be "strictly necessary."

The Digital Omnibus Proposal is relaxing the standards. The Proposal argues that since correcting biases is of paramount importance, the processing of this data should be made easier. Consequently, it proposes to relax the "strictly necessary" requirement to a mere "necessary" standard.

Article 9 of the GDPR is designed to protect the most intimate aspects of a person's identity through a general prohibition. Any derogation from such a fundamental prohibition must be interpreted as narrowly and strictly as possible. By removing the word "strictly," the proposal lowers the bar, potentially leading to the excessive processing of sensitive data. This shift contradicts the core logic of the GDPR and weakens the safeguards that protect individuals from intrusive data processing.