DE-ANONYMIZATION OF PERSONAL DATA IN THE PROCESS OF USING AI MODELS. ISSUES OF RESPONSIBILITY IN THE AI VALUE CHAIN.

Irina Tsakova

International Business School, 14 Gurko St., Botevgrad, 2140, Bulgaria

Abstract

This article examines the issues and risks related to the de-anonymization of personal data and the responsibility across the value chain in the development and deployment of artificial intelligence (AI) systems. Four hypotheses of potential de-anonymization of personal data during the use of AI systems are explored:(1) when the goal and intended purpose of the AI system are preserved;(2) when the goal and intended purpose of the AI system are changed;(3) when the system's goal and purpose are preserved, but the processing purpose changes;(4) when the goal of the system and data processing remains the same, but de-anonymization occurs due to technical or behavioral evolution of the system .Each hypothesis is analyzed through the lens of the requirements of the General Data Protection Regulation (GDPR, EU 2016), the Artificial Intelligence Act (EU 2024), and Opinion 28/2024 (EDPB, 2024). Based on the analysis of four regulatory and technical scenarios related to the purpose, intended use, and behavior of AI systems, the article discusses technological vulnerabilities such as behavioral identification, automated self-learning, and correlation with external sources. The goal is to clarify the conditions under which de-anonymization can occur despite formal compliance, and to specify the responsibility of actors in the supplier–implementer chain, including in cases of autonomous adaptation of AI models.

Keywords: *anonymisation, de-anonymisation, artificial intelligence, personal data*

INTRODUCTION

The Artificial Intelligence Act (AI Act, EU 2024) introduces a multilayered regulatory framework in which AI systems are viewed through the lens of the risk they pose—from unacceptable and high to limited. Each AI system is associated with specific obligations for the actors in the "value chain." Anonymization of personal data holds particular importance in this context, as improperly applied or outdated techniques can lead to re-identification. The Article 29 Working Party's Opinion 05/2014 on data anonymization techniques outlines classical anonymization methods such as generalization and randomization (Emam et al., 2014), but these approaches are often insufficient in the AI context. The reason lies in the capabilities of modern AI models to "memorize" training data, identify hidden dependencies, or even reproduce real-world information, which significantly increases the risk of data de-anonymization. Beyond traditional risks such as re-identification, record linkage, and extraction of sensitive information, new threats emerge in the AI environment—like inference, inversion, and extraction attacks, which are practically infeasible using traditional data analysis methods. To address these, the EDPB recommends using modern approaches—differential privacy, federated learning, homomorphic encryption, and synthetic data (EDPB, 2022). However, significant regulatory and methodological gaps remain. There are no clear criteria for determining when data can be considered anonymized in the context of AI, no certification or audit mechanisms for applied techniques, and the risk assessment approach is often inconsistent. Moreover, current regulations do not differentiate between sensitivity levels of various AI models - linguistic, visual, behavioral - which hampers the adaptation of protective measures to specific technologies.

These observations highlight the need for more precise regulation that considers the evolving nature of AI systems and the risk of privacy breaches. The introduction of minimum standards, re-identification testing, and periodic re-evaluation of anonymization emerge as necessary steps toward sustainable personal data protection in the age of artificial intelligence.

1. LITERATURE REVIEW

Modern academic literature highlights that standard anonymization methods such as k-anonymity, l-diversity, and even differential privacy are not fully effective in the context of artificial intelligence (AI) systems, particularly due to high dimensionality, correlations among data, and the need to maintain high model accuracy. According to Prajapati et al. (2024), classical approaches like k-anonymity and l-diversity prove ineffective in AI systems because of the dense information content and the necessity for precise models. They emphasize the trade-off between privacy and accuracy when using differential privacy and generative models. This aligns with the prevailing view in the literature that full anonymization is not always possible without losing data utility (Hafsa & Servani, 2024; Yiping et al., 2024). Aufschläger et al. (2023) also note that even anonymized tabular data can be de-anonymized by linking with external sources. Torres & Olivares (2024) demonstrate that quasi-identifiers allow for precise de-anonymization through combinations of attributes, and Deuber et al. (2023) show the ineffectiveness of classical techniques (e.g., name removal) when applied to court documents. Gadotti et al. (2024) describe anonymization as an "imperfect science," stating that modern techniques do not guarantee safety in complex AI scenarios. The main risks identified in the literature include: Linkage with external sources (linkage attacks), Recognizability through behavioral and linguistic models, Loss of data utility due to excessive anonymization, Algorithmic bias and unintentional manipulation (Bielova & Byelov, 2023). There is consensus on the need for new approaches that combine: Differential privacy (Prajapati et al., 2024; Gadotti et al., 2024), Use of synthetic data (Alabdulwahab et al., 2024; Gadotti et al., 2024), Federated learning (Gala et al., 2023; Prajapati et al., 2024). At the same time, differential privacy is considered the gold standard, but with limitations regarding utility (Qinyi et al., 2025; Ismat et al., 2022; Kaissis et al., 2023). On the other hand, synthetic data has the potential to reduce de-anonymization risks (Sarmin et al., 2024; Cairo, 2023; Gadotti et al., 2024). There are also differences in approaches. Some authors emphasize automation of anonymization through AI. Csányi et al. (2021) examine the dual role of AI in protecting and compromising privacy in data sharing. They stress that while AI can be a powerful tool for privacy protection via advanced anonymization techniques, it can also be used to extract sensitive information—even when data is supposedly protected. Others remain skeptical about effective anonymization in legal and sensitive contexts (Deuber et al., 2023). Furthermore, the study "AI-Driven Anonymization: Protecting Personal Data Privacy While Leveraging Machine Learning" (Yan et al., 2024) discusses anonymization challenges in the AI context. The authors highlight that even with the application of modern anonymization techniques, AI models can identify hidden patterns and dependencies in data, increasing the risk of re-identification. They also analyze various de-anonymization techniques, including inference and reconstruction attacks. From both legal and technical standpoints, it is believed that a balanced approach requires a combination of legislative mechanisms, technical safeguards, and ethical standards (Mamanazarov, 2024).

This literature supports the main thesis of this article - that anonymization in AI is not a trivial process and requires new methodological and regulatory approaches that go beyond classical paradigms.

2. ANALYSIS OF DE-ANONYMIZATION HYPOTHESES

2.1. Context

According to the AI Act, every AI system is developed with a specific purpose and intended use, which define its lawfulness and acceptability within the European Union. Article 3 defines "intended use" as the specific functionality for which the system is designed and employed, while "purpose" describes the context and tasks the system is supposed to perform. Moreover, the purposes of an AI system may differ from its intended use in a specific context (Recital 12). In addition, Article 4 requires providers and deployers to ensure that the system is used only for the originally defined purpose, and any deviation may result in a breach of the regulatory framework. In this context, Opinion 28/2024 provides specific guidance on the processing of personal data during the development and use of AI models (EDPB, 2024). The Opinion addresses the legal significance of the training phase, the residual risks of de-anonymization, and the responsibilities of various actors in processing, including providers and deployers. It is particularly emphasized that if a trained AI model reproduces or outputs data that can be

attributed to the same specific individuals whose data was used in the training process, then the system cannot be considered to be processing anonymized data. This is crucial for assessing compliance with GDPR and the AI Act. Recitals 67 and 69 of the AI Act emphasize that data administration practices must include transparency about the initial purpose, and data protection must be guaranteed throughout the lifecycle of the AI system. Additionally, Article 10, paragraphs 2 - 5 of the AI Act introduces requirements for the quality, representativeness, contextuality, and relevance of training and testing datasets in accordance with the system's intended use.

Therefore, it is essential not only to assess whether the training data were lawfully collected, but also whether the system, when in use, produces output that refers back to the same data subjects. If this occurs - even with an unchanged purpose - there is a real risk of de-anonymization, requiring reassessment of legality and the introduction of additional safeguards. This is what we consider below in each of the hypotheses.

2.2. Hypothesis 1: De-anonymization of Personal Data with Unchanged Purpose and Intended Use of the AI System

As noted by Prajapati et al. (2024), due to the high dimensionality and complex dependencies in the data used to train AI systems, the likelihood of de-anonymization remains high even when the system's original purpose and intended use are preserved. According to *Scenario 1* discussed in Opinion 28/2024 by the EDPB, if personal data was unlawfully processed during training and remains embedded in the model, any subsequent use by the same controller may be considered unlawful. This is particularly relevant for systems that allow regurgitation or extraction of personal data through query attacks, such as large language models. The Opinion explicitly states that when assessing anonymity, the possibility of direct or probabilistic identification based on current and future technological means must be considered. Therefore, if during the use of the system it becomes possible to recover the identity of data subjects via real-world queries, the model cannot be deemed anonymous. Responsibility lies with both the provider and the deployer, especially if they fail to take adequate measures to minimize or prove anonymity. In this scenario, data becomes de-anonymized during the use of the AI system without any change in its goal or purpose. The main risks here may include insufficient anonymisation of the original data - using techniques such as pseudonymisation, which does not fully remove the risk of reverse identification and may lead to re-identification through linkage attacks between anonymised and public data, inference attacks to extract sensitive information (inference attacks), reconstruction of original data using machine learning models (reconstruction attacks), and using combinations of metadata to identify individuals (attribute linkage attacks) (Yan et al., 2024), data correlation with external databases - combining anonymised records with open databases, e.g. combining health records with social network data (Wong et al., 2024). These risks are more likely to occur during the use phase, when the model processes real-world data that can be linked to external sources. However, the risk can also arise during model training, especially if the datasets used are not sufficiently well anonymized and contain latent dependencies that allow reconstruction of the subjects' identities.

In this hypothesis, the provider must ensure adequate anonymization measures, while the deployer bears responsibility for proper use. These measures may be proven through documentation of anonymization techniques, results of independent audits, compliance certificates with GDPR and the AI Act, and regular risk assessments. The deployer, in turn, may demonstrate proper use through data processing logs, internal control procedures, and evidence of staff training in data security. It is worth noting that Opinion 28/2024 stresses that if personal data is retained within an AI model, that model cannot be considered fully anonymous. The Opinion also highlights the need to assess re-identification risk, especially when anonymized data can be linked with external sources. According to *Scenario 2*, if personal data was unlawfully used in training by one controller and a different controller (deployer) uses the model later without verifying the legality of the original processing, the new controller may also be held liable. This underscores the need for independent assessments by deployers, including Data Protection Impact Assessments (DPIAs), checks on the origin of the data, and evaluation of residual re-identification risks (EDPB, Opinion 28/2024).

2.3. Hypothesis 2: De-anonymization due to a change in the purpose and intended use of the AI system

From the perspective of responsibility, it is important to clarify in advance that the hypothesis under consideration does not refer to the cases, according to Art. 25, (1), c) of the AI Act, in which, if the deployer modifies the intended use of an AI system, including a general-purpose AI system that is not classified as high-risk and has already been placed on the market or put into service, in such a way that the relevant AI system becomes a high-risk AI system, they shall be considered a provider, with all resulting obligations and consequences, including regarding the processing of personal data. This hypothesis considers cases in which the AI system is used for a new, unforeseen purpose, without the system becoming high-risk and outside the scope of "reasonably foreseeable misuse" (Art. 3, (13) of the AI Act), which leads to the de-anonymization of the data.

It is possible that the deployer consciously changes the purpose of the system to reprocess the data for a different aim, which could lead to the de-anonymization of the data. An example of such reprocessing in a new context is a system originally intended for analyzing purchasing habits for optimizing marketing campaigns, being used to create detailed user profiles that allow direct identification of individuals. This can be done by combining the purchase data with other sources, such as data from social media or mobile applications, leading to de-anonymization of users (Torres & Olivares, 2024). In this case, the deployer may intentionally change the processing purpose, using the data to create credit ratings or for targeted advertising, which exceeds the initial intended use of the system. Even if direct identifiers are removed, the presence of quasi-identifiers – such as age, location, and behavioral patterns – allows for a high degree of de-anonymization, especially when the data is used in new contexts (Torres & Olivares, 2024; Aufschläger et al., 2023). This risk increases with a change in the intended use of the AI system, which matches with the situation discussed in the current hypothesis. In this case, responsibility should lie with the deployer who changed the purpose or intended use of the AI system, as a result of which the data were de-anonymized, and both conditions must be applied cumulatively, with a causal link between the change in the purpose or intended use of the AI system and the de-anonymization of the data. It is also important whether the deployer exercises control over incoming data within the meaning of Art. 26, (4) of the AI Act, as well as whether the provider has ensured the necessary technical and organizational measures to prevent unlawful change of the original purpose of the system. In the latter case, the provider's responsibility may be engaged if sufficient restrictions have not been implemented.

2.4. Hypothesis 3: De-anonymization of the data with preserved system purpose and intended use, but change in the purpose of processing

Aufschläger et al. (2023) emphasize that even anonymized data, shared in tabular or aggregated form, may later lead to de-anonymization, especially when combined with external sources. This risk is particularly relevant in cases of data reuse where the original purpose for collection has changed or no longer aligns with the new intended use. According to Art. 6 (4) of the GDPR, in the event of a change in the processing purpose, an assessment must be conducted to determine whether the new purpose is compatible with the original one, based on the minimum required criteria. In Opinion 28/2024 by the EDPB, it is stressed that controllers must also consider the "reasonable expectations of data subjects". If collected data is used for incompatible or unexpected purposes—such as commercial, marketing, or profiling objectives—this may violate the principles of transparency and predictability. In the hypothesis under consideration, the AI system continues to be used according to its originally defined functionality and intended use, but there is a change in the purpose for which the personal data is processed. This means that even though the system itself is not reprogrammed or used outside its defined scope, the data collected through it is processed for a different, previously undeclared or incompatible purpose. Changing the processing purpose, even without technical modification of the system, presents a significant risk of de-anonymization, as the new context may enable re-identification through data linkage, especially if there are no additional protective measures. For example, processing health data originally collected for research purposes for commercial use, such as marketing or insurance. In this case, the deployer uses the output data for a different analytical purpose without altering the system's technical parameters. In this sense, Opinion 5/2022 of the EDPB warns that any contextual change in processing should be considered in light of the de-anonymization risk and the need to reassess the legal

basis. The responsibility for data protection in this case lies with the deployer, who acts as the data controller under the GDPR and makes decisions regarding the purpose and means of data processing. The deployer is obliged to ensure that personal data is not used for purposes that are incompatible with the original legal framework of the deployed system or that compromise, through de-anonymization, data used during system training. The deployer should conduct periodic Data Protection Impact Assessments (DPIAs) and maintain documentation of all processing purposes. If the provider has not enabled traceability and restriction of undesired deviations from permitted purposes, their responsibility cannot be entirely excluded in this case either.

2.5. Hypothesis 4: De-anonymization with preserved system purpose and data processing purpose

In this scenario, the AI system is used entirely in accordance with its originally declared purpose and intended use, and the data processing purpose has not been changed. Nevertheless, de-anonymization occurs due to the improvement of the system's analytical capabilities or the evolution of the technology. This risk is typical for systems that are continuously trained, adapt their behavior, or collect new data from the environment.

Anonymization in the context of adaptive AI systems is an uncertain process because the dynamic nature of self-learning models enables the restoration of identities through correlations and behavioral dependencies (Gadotti et al., 2024). This conclusion confirms the thesis that even with preserved system and processing purposes, the risk of de-anonymization remains real and requires additional safeguards. According to *Scenario 3* from Opinion 28/2024 of the EDPB, when initially there was unlawful processing of personal data during training, but later the model is completely anonymous, subsequent processing carried out with new data within the system may be permissible. Nevertheless, if de-anonymization occurs during use as a result of the system's adaptive behavior, it is considered a new event with new responsibility. The Opinion explicitly recommends that supervisory authorities check not only the initial legality of the data but also whether the actual behavior of the system corresponds to the declared anonymity. This recognizes that de-anonymization may occur even in formal compliance, if the technical design enables involuntary data disclosure through the autonomous actions of AI models. The autonomy and adaptability of AI systems significantly influence the ability to maintain effective anonymization. These systems can self-learn, adapt to new data and operating conditions, and make independent decisions based on accumulated experience without direct human intervention. This characteristic increases the risk of de-anonymization, as the system itself may begin to identify patterns and dependencies that lead to re-identification of subjects. As noted by Majeed & Hwang (2023), AI models may consciously or unconsciously exploit weaknesses in anonymization mechanisms, making them active participants in the de-anonymization process, rather than neutral tools. Here, risks of de-anonymization through linking anonymized data with publicly available sources should also be considered, as "smaller populations are more vulnerable to de-anonymization" (Wong et al., 2024).

It is the responsibility of providers to design systems in a way that prevents unforeseeable harmful effects, including de-anonymization. They must provide deployers with mechanisms to monitor the system’s behavior, as well as options to limit its autonomous development when this leads to deviations from processing objectives. In turn, deployers are responsible for monitoring and managing the system’s behavior in a real environment, and for taking corrective measures, including stopping the system, modifying algorithms, or conducting a new Data Protection Impact Assessment (DPIA) and making the necessary notifications. This means that both parties are obliged to foresee the possibility of re-identification even in seemingly anonymous processing, which requires a high level of control, transparency, and collaboration between providers and deployers.

3. CONCLUSIONS

De-anonymization of data when using AI systems is a complex and multidimensional issue that is becoming increasingly significant with the advent of more autonomous and adaptive technologies. Based on the reviewed possibilities for re-identification, the role of AI autonomy, as well as the responsibility of providers and deployers in the presented four hypotheses of de-anonymization, the following key conclusions can be drawn: De-anonymization may occur even when the AI system’s goal

and intended use remain unchanged, due to the application of inappropriate anonymization techniques by the provider or a poorly assessed risk—for example, the possibility of correlation with external data or the system's self-learning capabilities. Additionally, changing the purpose of the system or the purpose of personal data processing, without changing the goal of the AI system, carries a high risk of unlawful data processing, including the possibility of de-anonymization, which implies a higher level of responsibility for deployers. On the other hand, the autonomy and adaptability of AI increase the likelihood of unforeseen consequences, including re-identification, which requires stricter oversight and control. In this regard, providers and deployers should be equipped with appropriate tools to monitor and track the behavior of the AI system, including conducting tailored Data Protection Impact Assessments (DPIAs) that explicitly address de-anonymization risks across all stages - training, deployment, and operational use. To this end, various tools for risk analysis and visualization can be used to help identify vulnerabilities when publishing anonymized data (Adamakis et al., 2023), along with strict documentation of the goals and intended use of AI systems and transparency in communication with data subjects. Moreover, technical safeguards should be developed to control and, where necessary, restrict the autonomous behavior of AI in contexts where there is a risk of de-anonymization. Finally, the responsibilities along the provider–deployer chain should be clearly defined. The above also implies the need for innovative regulatory and technological mechanisms aimed at minimizing the risks of AI misuse in the context of personal data (Majeed et al., 2023).

REFERENCES

1. Aufschläger, P. et al. (2023). Data Anonymization and Re-identification in Structured Datasets. Journal of Privacy and Data Security. https://doi.org/10.1007/s10207-023-00601-3
2. Adamakis, E., Boch, M., Bampoulidis, A., Margetis, G., Gindl, S., & Stephanidis, C. (2023). DaRAV: A Tool for Visualizing De-Anonymization Risks. 2023 IEEE 39th International Conference on Data Engineering (ICDE), 3655-3658. https://doi.org/10.1109/ICDE55515.2023.00291.
3. Bielova, M., & Byelov, D. (2023). Challenges and threats of personal data protection in working with artificial intelligence. Uzhhorod National University Herald. Series: Law. https://doi.org/10.24144/2307-3322.2023.79.2.2.
4. Csányi, G.M.; Nagy, D.; Vági, R.; Vadász, J.P.; Orosz, T. Challenges and Open Problems of Legal Document Anonymization. Symmetry 2021, 13, 1490. https://doi.org/10.3390/sym13081490
5. Qinyi Liu, Ronas Shakya, Mohammad Khalil, and Jelena Jovanovic. 2025. Advancing privacy in learning analytics using differential privacy. In Proceedings of the 15th International Learning Analytics and Knowledge Conference (LAK '25). Association for Computing Machinery, New York, NY, USA, 181–191. https://doi.org/10.1145/3706468.3706493;
6. Jahan Sarmin, F., Rahman Sarkar, A., Wang, Y., & Mohammed, N. (2024). Synthetic Data: Revisiting the Privacy-Utility Trade-off. arXiv e-prints, arXiv-2407.
7. Deuber, D., Keuchen, M., & Christin, N. (2023). Assessing Anonymity Techniques Employed in German Court Decisions: A {De-Anonymization} Experiment. In 32nd USENIX Security Symposium (USENIX Security 23) (pp. 5199-5216 https://www.usenix.org/conference/usenixsecurity23/presentation/deuber;
8. Gadotti, A., Rocher, L., Houssiau, F., Creţu, A.-M., & de Montjoye, Y.-A. (2024). Anonymization: The Imperfect Science of Using Data While Preserving Privacy. Science Advances, 10(29), eadn7053. https://pubmed.ncbi.nlm.nih.gov/39018389/;
9. Ismat Jarin and Birhanu Eshete. 2022. DP-UTIL: Comprehensive Utility Analysis of Differential Privacy in Machine Learning. In Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy (CODASPY '22). Association for Computing Machinery, New York, NY, USA, 41–52. https://doi.org/10.1145/3508398.3511513

10. Alabdulwahab, S., Kim, Y.-T., & Son, Y. (2024). Privacy-Preserving Synthetic Data Generation Method for IoT-Sensor Network IDS Using CTGAN. Sensors, 24(22), 7389. https://doi.org/10.3390/s24227389
11. Gala, J., Gandhi, D., Mehta, J., & Talat, Z. (2023). A Federated Approach for Hate Speech Detection. Conference of the European Chapter of the Association for Computational Linguistics.
12. Hafsa, L., & Sevrani, K. (2024). Preserving privacy in medical images while still enabling AI-driven research: A comprehensive review. 2024 13th Mediterranean Conference on Embedded Computing (MECO), 1-5. https://doi.org/10.1109/MECO62516.2024.10577795.
13. A. Majeed and S. O. Hwang, "When AI Meets Information Privacy: The Adversarial Role of AI in Data Sharing Scenario," in IEEE Access, vol. 11, pp. 76177-76195, 2023, doi: 10.1109/ACCESS.2023.3297646
14. Mamanazarov, Sardor. (2024). De-identification and anonymization: legal and technical approaches. Tsul legal report. 5. 25-37. https://doi.org/10.51788/tsul.lr.5.1./TCYN1311;
15. Yang,L.,Tian,M.,Xin,D.,Cheng,Q.,Zheng,J. (2024). AI-driven anonymization: Protecting personal data privacy while leveraging machine learning. Applied and Computational Engineering, 67,273-279, https://doi.org/10.54254/2755-2721/67/2024MA0053
16. Prajapati, M., Upadhyay, A., Rezaie, M., & Dongradive, J. (2024). A comparative Study on AI-Driven Anonymization Techniques for Protecting Personal Data. International Journal For Multidisciplinary Research. https://doi.org/10.36948/ijfmr.2024.v06i04.24770;
17. Han, Yiping & Lu, Xinqian. (2024). Is Data Anonymization an Effective Way to Protect Privacy or Not. International Journal of Computer Science and Information Technology. 4. 152-156. 10.62051/ijcsit.v4n3.15;
18. Cairo, Michael (2023) "Synthetic Data and GDPR Compliance: How Artifical Intelligence Might Resolve the Privacy-Utility Tradeoff," Journal of Technology Law & Policy: Vol. 28: Iss. 1, Article 4.Available at: https://scholarship.law.ufl.edu/jtlp/vol28/iss1/4
19. W. Wong, Z. Alomari, Y. Liu and L. Jura, "Linkage Deanonymization Risks, Data-Matching and Privacy: A Case Study," 2024 8th International Conference on Cryptography, Security and Privacy (CSP), Osaka, Japan, 2024, pp. 10-16, doi: 10.1109/CSP62567.2024.00010
20. Emam, K. & Alvarez, C.. (2014). A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques. International Data Privacy Law. 5. 73-87. https://doi.org/10.1093/idpl/ipu033 ;
21. Georgios Kaissis, Alexander Ziller, Stefan Kolek, Anneliese Riess, and Daniel Rueckert. 2023. Optimal privacy guarantees for a relaxed threat model: Addressing sub-optimal adversaries in differentially private machine learning. 37th Conference on Neural Information Processing Systems (NeurIPS 2023). (2023). https://proceedings.neurips.cc/paper_files/paper/2023/file/aea831d6c7af37fd4230937225be3414-Paper-Conference.pd
22. Torres, N., & Olivares, P. (2024). De-Anonymizing Users across Rating Datasets via Record Linkage and Quasi-Identifier Attacks. Data, 9, 75. https://doi.org/10.3390/data9060075;
23. Regulation (EU) 2024/1689 of the European parliament and of the Council of 13 june 2024, laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) no 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial intelligence act);
24. Regulation (EU) 2016/679 of the European parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC;

25. European Data Protection Board (EDPB), Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models;
26. Article 29 Working Party, Opinion 05/2014 on data anonymization techniques;
27. European Data Protection Board (EDPB), Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement;