

Harnessing AI for AML/CFT: Legal grounds for training AI on personal data for AML/CFT under EU data protection law

Authors	Roussos,Manos; Hajduk,Pawel
Published in	Information & Communications Technology Law
DOI	10.1080/13600834.2025.2510748
Publication Date	2026
Document Version	publishersversion
Link	https://research.tilburguniversity.edu/en/publications/dbf233fc-9329-4bb5-a433-f3c1a0dc705e
Citation	Roussos, M & Hajduk, P 2026, 'Harnessing AI for AML/CFT : Legal grounds for training AI on personal data for AML/CFT under EU data protection law', Information & Communications Technology Law, vol. 35, no. 1, pp. 72-92. https://doi.org/10.1080/13600834.2025.2510748
Download Date	2026-05-17 12:39:53
Rights	<p>General rights</p> <p>Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.</p> <ul style="list-style-type: none"> - Users may download and print one copy of any publication from the public portal for the purpose of private study or research. - You may not further distribute the material or use it for any profit-making activity or commercial gain - You may freely distribute the URL identifying the publication in the public portal" <p>Take down policy</p> <p>If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.</p>



Harnessing AI for AML/CFT: Legal Grounds for Training AI on Personal Data for AML/CFT under EU Data Protection Law

Manos Roussos & Paweł Hajduk

To cite this article: Manos Roussos & Paweł Hajduk (2026) Harnessing AI for AML/CFT: Legal Grounds for Training AI on Personal Data for AML/CFT under EU Data Protection Law, *Information & Communications Technology Law*, 35:1, 72-92, DOI: [10.1080/13600834.2025.2510748](https://doi.org/10.1080/13600834.2025.2510748)

To link to this article: <https://doi.org/10.1080/13600834.2025.2510748>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 31 May 2025.



Submit your article to this journal [↗](#)



Article views: 2029



View related articles [↗](#)



View Crossmark data [↗](#)

Harnessing AI for AML/CFT: Legal Grounds for Training AI on Personal Data for AML/CFT under EU Data Protection Law

Manos Roussos^a and Paweł Hajduk^b

^aTilburg Institute for Law, Technology, and Society (TILT), Tilburg, The Netherlands; ^bDepartment of Informatics Law, Cardinal Stefan Wyszyński University, Warsaw, Poland

ABSTRACT

AI systems can assist in fulfilling AML/CFT obligations within the revised EU AML framework. To function accurately, these AI-enhanced AML systems require extensive training on datasets, including personal data. This paper examines the legal grounds under the General Data Protection Regulation (GDPR) for processing such data, with a focus on compliance with legal obligations [Article 6(1)(c) GDPR] and legitimate interest [Article 6(1)(f) GDPR]. The paper argues that, while legal obligation may not provide a sufficient basis due to the lack of explicit mandates requiring AI use, legitimate interest presents a viable alternative, dependent on a rigorous test. By scrutinising the necessity of balancing financial institutions' need for AI-enhanced AML/CFT tools with EU data protection law, this paper underscores the significance of safeguards to mitigate risks associated with such tools, including bias, transparency shortcomings, and challenges in exercising data subject rights.

KEYWORDS

Anti-money laundering; AML; CFT; AI; legal grounds; personal data processing

1. Introduction

In the foreseeable future, we may transition from democratically governed societies to an 'algocracy', where algorithms dictate the course of action across several sectors and different aspects of life. The financial industry is a prime illustration of this ongoing transformation, as AI systems¹ have revolutionised this sector with applications ranging from customer interaction to fraud detection and anti-money laundering efforts.² While the adoption of AI in finance is undeniably accelerating, promising enhanced efficiency

CONTACT Manos Roussos  E.Roussos@tilburguniversity.edu

¹EU law defines 'AI system' in 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) [2024] OJ L 1689 as

a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

²Charl Maree and others, 'Towards Responsible AI for Financial Transactions' (2020) IEEE Symposium Series on Computational Intelligence (SSCI) 16.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and regulatory compliance, the precise impact of AI on the financial sector remains uncertain. It necessitates careful consideration of future regulatory responses.³

Although the financial services industry has a rich history of regulatory practices dating back centuries (for example, trading regulations prohibiting short selling arose in the seventeenth-century Netherlands),⁴ introducing AI into financial systems presents a new set of challenges. Relying on AI systems to make informed decisions requires training them on large datasets, which exposes many financial institutions to multiple risks, including vulnerability to manipulation, lack of transparency in decision-making, and other data protection concerns.⁵ Nevertheless, due to the rapid advancement of AI technology, various banks and other financial institutions are increasingly adopting it for fraud prevention, with recent reports indicating that 70% of these institutions are already using AI to fight financial crime, such as fraud.⁶

These fraud prevention strategies naturally align with Anti-Money Laundering (AML) efforts undertaken by obliged entities within the anti-money laundering and countering the financing of terrorism (AML/CFT) framework. These entities are regarded as gatekeepers due to their vital contribution in identifying potential illicit activities.⁷ The term '*obliged entities*' encompasses credit institutions, financial institutions and other natural or legal persons,⁸ all of which are required by law to implement specific measures aimed at preventing money laundering;⁹ in the context of this research, AML efforts made by financial institutions are particularly relevant.

Money laundering is defined as

the process by which criminals "clean" the benefits of their activities to hide their illegal origin. It is usually associated with the types of organised crime that generate huge profits in cash, such as trafficking in drugs, weapons and human beings as well as fraud.¹⁰

AML refers to the '*the set of laws, regulations, and procedures intended to prevent criminals from disguising illegally accumulated funds as legitimate income*',¹¹ while the primary objectives of AML rules are '*to target activities like market manipulation, trade of illegal goods, terrorist financing, and corruption of public fund*'.¹² The history of detecting the

³Joost van der Burgt, *General Principles for the Use of Artificial Intelligence in the Financial Sector* (De Nederlandsche Bank (DNB), 2019) 5.

⁴Paul Matthias, 'Artificial Intelligence in Financial Services: New Risks and the Need for More Regulation' in Markus D Dubber, Frank Pasquale and Sunit Das (eds), *The Cambridge Handbook of Responsible Artificial Intelligence* (CUP, 2022) 364.

⁵El Bachir Boukheroua and Ghiath Shabsigh, 'Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance' (2021) IMF Department Papers Series 15.

⁶PYMNTS, 'Seven in 10 Financial Institutions Use AI and ML to Combat Fraud' (PYMNTS, 26 March 2024) <<https://www.pymnts.com/news/security-and-risk/2024/seven-in-10-financial-institutions-use-ai-and-ml-to-combat-fraud/>> accessed 6 May 2025.

⁷Council of the EU, 'Anti-money Laundering: Council and Parliament Strike Deal on Stricter Rules' *Press Release* (18 January 2024) <<https://www.consilium.europa.eu/en/press/press-releases/2024/01/18/anti-money-laundering-council-and-parliament-strike-deal-on-stricter-rules/>> accessed 6 May 2025.

⁸Regulation (EU) 2024/1624 of the European Parliament and of the Council of 31 May 2024 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing [2024] OJ L 1624, art. 3.

⁹Thomas Wahl, 'The EU's New AML Single Rulebook Regulation' *Eucrim* (18 July 2024) <<https://eucrim.eu/news/the-eu-new-aml-single-rulebook-regulation/>> accessed 6 May 2025.

¹⁰European Commission, 'Money laundering' (Migration and Home Affairs) <https://home-affairs.ec.europa.eu/policies/internal-security/organised-crime-and-human-trafficking/money-laundering_en> accessed 6 May 2025.

¹¹Shant Vosgueritchian, 'A Helping Hand: How Artificial Intelligence Can Help Financial Institutions Comply with AML/BSA Requirements' (2019) 39 *Review of Banking & Financial Law* 181.

¹²*Ibid* 182.

crime of money laundering, which can be defined as an act of '*disguising illegal or tax-avoided money and bringing it into formal monetary channels*',¹³ dates back to the 1970s when financial institutions initiated reporting large transactions to authorities. Towards the end of the 1990s, statistical approaches like Bayesian models and temporal sequence matching emerged for AML pattern detection. By 2004, the application of machine learning methods started gaining traction for the same purpose.¹⁴ Machine learning techniques are a subset of AI, based on training computer systems to obtain insights from data, recognise patterns, and autonomously make decisions with little human input.¹⁵ Despite the multitude of available AI techniques, machine learning stands out as the most relevant in the context of AML.¹⁶

The 2018 Danske Bank scandal was a pivotal event influencing the adoption of AI systems for detecting and flagging suspicious activity.¹⁷ Characterised as one of the largest money laundering scandals in European history, the case involves €200 billion in suspicious transactions from the Estonian branch of the bank between 2007 and 2016.¹⁸ Shortcomings in traditional monitoring methods were exposed due to this scandal, leading to an increased adoption of AI systems to identify and flag suspicious activities. Danske Bank themselves, for instance, decided to implement AI systems designed to accelerate decision-making and ensure timely reporting to regulatory organisations.¹⁹ Similarly, other banks have adopted AI-based tools; Standard Chartered Bank employs AI-powered systems to monitor transactions in real-time, allowing suspicious activity to be flagged immediately and reducing opportunities for money laundering.²⁰ Another systemic bank broadly using AI is Citibank, which employs AI tools that analyse customer transactions and behavioural patterns across its global accounts, enabling immediate action when necessary.²¹ In general, the vast majority of banks worldwide employs some kind of AI for their works, either for payments, fraud detection, or just customer relations.²² AI has been used by banks, but notably, the first time that a specific legal provision was attributed to AI use by banks for AML/CFT purposes was only recently. More specifically, in May 2024, the EU published a new legislative package, aiming to clarify the European AML/CFT landscape; one of its components is the so-called AML Regulation (AMLR),²³ whose Article 76(5) provides for the use of processes involving AI by obliged entities.

¹³Abhishek Gupta and others, 'Overview of Money Laundering' in *Artificial Intelligence Applications in Banking and Financial Services, Future of Business and Finance* (Springer, July 2023) 1.

¹⁴Alhanouf Abdulrahman and others, 'Anti-money Laundering Systems: A Systematic Literature Review' (2020) 23 *Journal of Money Laundering Control* 834.

¹⁵The Financial Action Task Force (FATF), *Opportunities and Challenges of New Technologies for AML/CTF* (Paris, July 2021) 22.

¹⁶Jingguang Han and others, 'Artificial Intelligence for Anti-money Laundering: A Review and Extension' (2020) 2 *Digital Finance* 211–39.

¹⁷Pythagoras Blog Team, 'Danske Bank Scandal: Turning Point AML Compliance of Banking Sector' *Pythagoras* (10 April 2024) <<https://pythagoras-solutions.com/en/insights/danske-bank-scandal-turning-point-aml-compliance-of-banking-sector>> accessed 6 May 2025.

¹⁸Elisabetta Bjerregaard and Tom Kirchmaier, *The Danske Bank Money Laundering Scandal: A Case Study* (Copenhagen Business School, September 2019) 4.

¹⁹AML Watcher, '7 Use Cases of Artificial Intelligence in Anti-Money Laundering' *AML Watcher* (2024) <<https://amlwatcher.com/blog/7-use-cases-of-artificial-intelligence-in-anti-money-laundering/>> accessed 6 May 2025.

²⁰*Ibid.*

²¹*Ibid.*

²²OECD, 'Regulatory approaches to Artificial Intelligence in finance' (OECD Artificial Intelligence Papers, No. 24, 2024) 15–16

²³Regulation (EU) 2024/1624 of the European Parliament and of the Council of 31 May 2024 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing

Accordingly, rather than directly identifying criminal behaviour, AI-powered AML systems may identify any abnormal transaction patterns made by customers, facilitating quicker regulatory interventions.²⁴ Yet, this approach can lead to potential violations of data protection regulations. Although customers might wish to understand why a transaction was flagged,²⁵ decisions of AI systems are not always transparent or readily justifiable. The algorithms behind AI tools, commonly called ‘black box’ algorithms, are difficult to explain, due to their potential unpredictability and autonomous ability to uncover patterns in the data they process, an issue often defined as the ‘black box effect’.²⁶ Distinct from traditional statistical methods, where humans select variables, AI algorithms do not directly link variables to outcomes, complicating the justification of their final results.²⁷ ‘Black box’ features of AI systems raise significant compliance concerns,²⁸ compelling governing bodies and financial institutions to address data protection considerations by ensuring that AI-driven decisions maintain fairness, transparency and accountability.²⁹

In this AI-driven landscape, data protection faces significant challenges as AI systems grow increasingly complex for the purposes of mass data collection and analysis. To function effectively and achieve the intended outcomes, such AI systems must be trained on data, including standard personal data (such as names, contact details and identification numbers of each customer) but also details about transactions, fund movements, and specific patterns in payments and money transfers. Determining the lawfulness of personal data collection is challenging.³⁰ Therefore, the main question discussed in this paper is what legal basis can be used to process personal data for training such AI systems. In that sense, ensuring compliance may necessitate strict controls to safeguard sensitive financial information.³¹ Some would argue that the lack of a detailed legal framework for adopting AI technology in financial services highlights the need for more precise guidance rather than additional rules; proponents of this idea call on regulators to evaluate existing legislation, including data protection regulations, to determine their relevance to AI systems.³² This is particularly relevant in light of the 2022 Dutch Court decision (*Bunq v. DNB*),³³ which paved the way for banks to use AI in risk monitoring as

²⁴Doron Goldbarsht, ‘Leveraging AI to Mitigate Money Laundering Risks in the Banking System’ in Z Bednarz and M Zalnieriute (eds), *Money, Power, and AI: Automated Banks and Automated States* (Cambridge University Press 2023) 58.

²⁵Bryce Goodman and Seth Flaxman, ‘European Union Regulations on Algorithmic Decision-making and a “Right to Explanation”’ (2017) 38 *AI Magazine* 55.

²⁶European Commission, ‘White Paper on Artificial Intelligence – A European Approach to Excellence and Trust’ (2020) 12.

²⁷Cary Coglianese, ‘Law and Empathy in the Automated State’ in Z Bednarz and M Zalnieriute (eds), *Money, Power, and AI: Automated Banks and Automated States* (Cambridge University Press 2023) 178.

²⁸Financial Stability Board, ‘Artificial Intelligence and Machine Learning in Financial Services’ (2017) 13.

²⁹Barry Quinn and others, ‘How Will Artificial Intelligence Affect Financial Regulation?’ *Economics Observatory* (18 October 2023) <<https://www.economicsobservatory.com/how-will-artificial-intelligence-affect-financial-regulation>> accessed 6 May 2025.

³⁰Tamara Tanaskovic, ‘The Regulatory Burdens of AI in AML Investigations’ *Medium* (5 November 2019) <<https://medium.com/@XXXtech/the-regulatory-burdens-of-ai-in-aml-investigations-afceca76546b>> accessed 6 May 2025.

³¹Barry Quinn and others, ‘How Will Artificial Intelligence Affect Financial Regulation?’ *Economics Observatory* (18 October 2023) <<https://www.economicsobservatory.com/how-will-artificial-intelligence-affect-financial-regulation>> accessed 6 May 2025.

³²Simon Toms and others, ‘How Regulators Worldwide Are Addressing the Adoption of AI in Financial Services’ *Skadden* (12 December 2023) <<https://www.skadden.com/insights/publications/2023/12/how-regulators-worldwide-are-addressing-the-adoption-of-ai-in-financial-services>> accessed 6 May 2025.

³³*Bunq BV v DNB*, C-21/323 and 21/1108, ECLI:NL:CBB:2022:707 (Trade and Industry Appeals Tribunal, 18 October 2022).

part of their AML strategies, a development that may necessitate extensive personal data processing to identify such risks effectively.³⁴

In terms of limitations, this paper only touches upon the issues related to the intersection between the updated 2024 EU AML/CFT framework, AI-enhanced tools, and data protection. For the purposes of this paper, we have not examined special categories of personal data, since these are not included in the list of Article 22 AMLR, which determines the personal data necessary for the Customer Due Diligence process.³⁵ In the context of AML/CFT, the only personal data of a sensitive nature that might be processed are often those relating to criminal proceedings and convictions. Processing of such personal data shall be allowed only when appropriate safeguards are established in law.³⁶ Besides, processing special categories of personal data as defined in Article 9 GDPR is not generally conceived as a customary or standard procedure for AML/CFT procedures, so the training of AI systems for AML/CFT purposes would not *in principio* be based on sensitive personal data. The analysis will be conducted in conjunction only with the 2024 EU AML/CFT framework, with the provisions of one of its legislative tools taken especially into account for the purposes of this paper: the AML Regulation.³⁷

2. AI systems and data protection risks

With great hopes attached to AI, new risks also emerge. This section identifies selected characteristics of AI that give rise to personal data protection risks and may affect the legal assessment of the bases for processing personal data. These risks include the transparency and explainability of AI, the accuracy of processing, and risks to individuals' reputations. There is also the risk of re-identifying anonymised data and challenges to the possibility of erasing data once entered into AI systems, generating a risk for rights to opt out.

The risks associated with AI relate to ensuring AI systems' transparency. The principle of transparency³⁸ [Article 5(1)(a) GDPR] is aligned with the principles of lawfulness and fairness in the GDPR.³⁹ In AI, a key transparency issue is the ambiguity of how the algorithms work, i.e. how they learn from the input and generate an output.⁴⁰ This is often referred to as the '*black-box*' phenomenon because it is difficult to reconstruct precisely how the algorithm works, including specifying which data from a dataset was used to generate

³⁴Moody's, 'Rise of the Machines – What Landmark Dutch AI Ruling Means for AML Compliance' *Moody's Blog* (22 November 2022) <<https://www.moody's.com/web/en/us/kyc/resources/insights/rise-machines-what-landmark-dutch-ai-ruling-means-aml-compliance.html>> accessed 6 May 2025.

³⁵These are the name, place and date of birth, nationality, place of residence and tax identification number of each individual concerned.

³⁶Eleni Kosta, *Report on the Implications for Data Protection of Mechanisms for Inter-State Exchanges of Data for Anti-Money Laundering/Countering Financing of Terrorism, and Tax Purposes* (Council of Europe 2021) <<https://rm.coe.int/t-pd-2021-4rev-inter-state-exchanges-of-data-for-tax-purposes-and-cml-1680a3ed30>> accessed 7 May 2025.

³⁷Regulation (EU) 2024/1624 of the European Parliament and of the Council of 31 May 2024 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing [2024] OJ L 1624.

³⁸Stefan Larsson and Fredrik Heintz, 'Transparency in Artificial Intelligence' [2020] 9(2) *Internet Policy Review* <<https://policyreview.info>> accessed 30 October 2024.

³⁹Laurens Naudts and others, 'Meaningful Transparency through Data Rights: A Multidimensional Analysis' [2022] 2022(37) *Amsterdam Law School Research Paper* accessed 30 October 2024.

⁴⁰Heike Felzmann and others, 'Towards Transparency by Design for Artificial Intelligence' [2020] 26(6) *Science and Engineering Ethics* 3333–61, 3334–35.

a given output,⁴¹ which translates into problems in explaining these operations to data subjects.⁴²

There are also risks associated with how the algorithms work regarding the correctness of AI output, including the fact that they may result in reputational damage to individuals. The principle of accuracy [Article 5(1)(d) GDPR] of personal data processing means that controllers are obliged to process data correctly and truthfully, taking ‘every reasonable step’.⁴³ Its infringement might potentially manifest in two forms. The so-called ‘AI hallucinations’ pertain to situations in which the outcome of the AI is not correct but is presented as a true statement.⁴⁴ DPAs have discussed this issue in the context of ChatGPT, highlighting a threat to the accuracy principle,⁴⁵ while OpenAI maintains that it is impossible to ensure complete correctness of outputs due to technical limitations.⁴⁶ Following these developments, recent literature suggests that the traditional understanding of the accuracy principle, advanced in the context of the correctness of data entries in controllers’ databases, may not be sufficient regarding AI systems.⁴⁷ These systems generate results based on probabilistic calculations and cannot always guarantee correctness. Thus, rather than adopting an absolutist stance, a more flexible approach is being considered, focusing on the introduction of safeguards to mitigate the risk of incorrectness.⁴⁸ While not making a decisive statement on this issue, it is vital to consider this circumstance when assessing the lawfulness of processing, particularly under the legitimate interest basis, as it may negatively impact the rights and freedoms of individuals.

The second issue relates to potentially discriminatory outputs. AI systems are trained on large datasets that may contain unfounded presuppositions about the role and position of, e.g. ethnic groups and nationalities.⁴⁹ These biased judgements, pre-existing in the datasets, are embedded into the AI systems. As a result, they produce discriminatory

⁴¹However, it should be noted that technical attempts are being made to minimise the risks for transparency, which is known under the umbrella term of ‘explainable AI’. Cf A Rai, ‘Explainable AI: From Black Box to Glass Box’ [2019] 48(1) *Journal of the Academy of Marketing Science* 138 <<https://doi.org/10.1007/s11747-019-00710-5>>.

⁴²Lilian Mitrou, ‘Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) “Artificial Intelligence-Proof?”’ [2018] *Information Systems Legislation & Regulations eJournal* 54–58 <<https://doi.org/10.2139/ssrn.3386914>>.

⁴³This is limited by the content of the principle. Article 5(1)(d) GDPR: ‘every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay’.

⁴⁴Ji Ziwei and others, ‘Survey of Hallucination in Natural Language Generation’ [2023] 55(12) *ACM Computing Surveys* <<https://doi.org/10.1145/3571730>> accessed 30 October 2024.

⁴⁵However, there is a need for a distinction between accuracy principle and statistical accuracy, Information Commissioner’s Office, ‘What Do We Need to Know about Accuracy and Statistical Accuracy?’ *ICO* (15 March 2023) <<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/what-do-we-need-to-know-about-accuracy-and-statistical-accuracy/>> accessed 4 January 2025.

⁴⁶Janos Meszaros and others, ‘Chatgpt: How Many Data Protection Principles Do You Comply with?’ [2023] 17–19 <https://kuleuven.limo.libis.be/discovery/fulldisplay?docid=lirias4127916&context=SearchWebhook&vid=32KUL_KUL:Lirias&lang=en&search_scope=lirias_profile&adaptor=SearchWebhook&tab=LIRIAS&query=any,contains,LIRIAS4127916&offset=0> accessed 30 October 2024. This preprint was available from November 2023 to early January 2024 on SSRN but was removed from SSRN. Its analysis was done prior to its removal.

⁴⁷Theodore Christakis, ‘AI Hallucinations and Data Subject Rights under the GDPR: Regulatory Perspectives and Industry Responses’ [2025] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5042191> accessed 6 May 2025.

⁴⁸Ibid.

⁴⁹F Zuiderveen Borgesius, *Discrimination, Artificial Intelligence, and Algorithmic Decision-making* (Council of Europe, Directorate General of Democracy, 2018) <<https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-%20making/1680925d73>> accessed 30 October 2024; Alexander Tischbirek, ‘Artificial Intelligence and Discrimination: Discriminating against Discriminatory Systems’ in Thomas Wischmeyer and Timo Rademacher (eds), *Regulating Artificial Intelligence* (Springer Nature, 2019); Philipp Hacker, ‘Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under EU Law’ [2018] 55(4) *Common Market Law Review* 1143–85.

outputs reflecting this imprinted prejudice.⁵⁰ Although discrimination does not always lead to a data protection infringement, it can affect it if it enables AI hallucinations, reinforcing the risk of incorrect outcomes.

Further, due to its ability to search for patterns,⁵¹ the use of AI can potentially lead to the re-identification of previously anonymised personal data. Anonymisation processes are not always conducted correctly and with sufficient depth.⁵² Correlating anonymised data with other data obtained by AI may pose a risk of re-identification; such identification may also occur accidentally.⁵³ There are instances demonstrating the possibility of the AI's ability to correlate different information, such as a prompt in which the word 'poem' was asked to be continuously generated, and the chatbot (in this case specifically the ChatGPT, i.e. generative AI), for an unknown reason, generated contact details for an individual.⁵⁴ The EDPB argues that AI systems might be considered anonymous. However, the threshold is high, depending on the technical solutions during the AI training, including data preparation and methods implemented to enhance privacy.⁵⁵ Data that has been previously properly anonymised does not fall under the GDPR,⁵⁶ but if re-identified, it becomes personal data again.⁵⁷

Moreover, the literature highlights the difficulty of deleting data once it has been entered into AI systems. Hence, it is doubtful whether an attempt⁵⁸ to delete data can be compelling, given that deleting data at the entry point cannot easily change how algorithms have already learned from data.⁵⁹ This calls into question the ability to effectively exercise the right to withdraw consent for processing data used to train the algorithm model (if this basis was chosen),⁷⁸ as well as the right to be forgotten (Article 17 GDPR), the right to rectification (Article 16 GDPR); and the right to object, particularly relevant particularly to the legitimate interest basis (Article 21 GDPR).

The legal assessment of these characteristics should be made in the context of the different stages of personal data processing within AI system development. The discussion paper 'Legal basis in data protection in the use of artificial intelligence' by the

⁵⁰Daniel Varona and JuanLuis Suárez, 'Discrimination, Bias, Fairness, and Trustworthy AI' [2022] 12(12) Applied Sciences <<https://doi.org/10.3390/app12125826>> accessed 30 October 2024.

⁵¹Emily M Weitzenboeck and others, 'The GDPR and Unstructured Data: Is Anonymization Possible?' [2022] 12(3) International Data Privacy Law <<https://doi.org/10.1093/idpl/ipac008>> accessed 30 October 2024 203.

⁵²J Willemson, 'Fifty Shades of Personal Data – Partial Re-identification and GDPR' in *Privacy Technologies and Policy, APF 2022*, Lecture Notes in Computer Science (Springer, 2022); Elizabeth A Brasher, 'Addressing the Failure of Anonymization: Guidance from the European Union's General Data Protection Regulation' [2018] 1 Columbia Business Law Review 209–53.

⁵³EDPB, 'Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models' EDPB (2024) <https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf> accessed 3 January 2025, paras 29–34, 35–43, especially para 38.

⁵⁴Tiernan Ray, 'ChatGPT can Leak Training Data, Violate Privacy, Says Google's DeepMind' ZDNET (4 December 2023) <https://www.zdnet.com/article/chatgpt-can-leak-source-data-violate-privacy-says-googles-deepmind/#google_vignette> accessed 27 February 2025.

⁵⁵Ibid 50–52.

⁵⁶Recital 26 GDPR.

⁵⁷F Di Cerbo and S Trabelsi, 'Towards Personal Data Identification and Anonymization Using Machine Learning Techniques' in A Benczúr and others (eds), *New Trends in Databases and Information Systems: ADBIS 2018. Communications in Computer and Information Science*, vol 909 (Springer, 2018) https://doi.org/10.1007/978-3-030-00063-9_13.

⁵⁸Antonio Ginart and others, 'Making AI Forget You: Data Deletion in Machine Learning' [2019] 32 Advances in Neural Information Processing Systems <https://proceedings.neurips.cc/paper_files/paper/2019/file/cb79f8fa58b91d3af6c9c991f63962d3-Paper.pdf> accessed 30 October 2024.

⁵⁹R Chourasia and N Shah, 'Forget Unlearning: Towards True Data-deletion in Machine Learning' [2023] 202 Proceedings of the 40th International Conference on Machine Learning, Proceedings of Machine Learning Research <<https://openreview.net/pdf/cfbfc2f43e5d5ac365020aa3db4f19384ddc3cae.pdf>> accessed 30 October 2024.

Data Protection Authority in Baden-Württemberg (Germany) distinguishes five phases of data protection processing within AI system development: collection of training data for AI; processing data for training AI; provision of AI applications; use of AI applications; and use of AI results.⁶⁰ This paper focuses on the training phase. The legal assessment should be conducted separately for each data processing phase. However, risks that affect the evaluation of individuals' rights and freedoms in one phase can also influence the legal assessment in another phase. This is because specific issues might not be apparent during the personal data collection phase but may emerge during training or when using AI applications. Despite being distinct phases, all these steps are functionally interconnected. Therefore, it can be argued that the outputs generated by AI systems based on collected and trained personal data can, on a case-by-case evaluation,⁶¹ affect the legal assessment of their processing throughout all phases.

3. The use of AI-Enhanced tools to improve AML/CFT processes

3.1. AI in AML: revolutionizing detection and compliance

To comply with AML/CFT laws, banks and other obliged entities must monitor transactions and report suspicious financial activities to competent authorities when reasonable grounds for money laundering-related concerns arise.⁶² The hefty penalties imposed for non-compliance with AML/CFT rules further incentivise financial institutions to adhere to AML regulations strictly.⁶³ To illustrate this, the total amount of fines for failing to comply with AML/CFT laws reached \$6.6 billion in 2023, a substantial increase from \$4.2 billion in 2022 and \$5.4 billion in 2021,⁶⁴ highlighting the growing regulatory scrutiny and escalating costs of non-compliance.

Nevertheless, banks often struggle to comply with AML measures due to the complexity of accurately identifying, assessing, and addressing money laundering risks, particularly during customer due diligence (CDD) processes.⁶⁵ According to Article 20 and Recital 51 of the newly established Anti-Money Laundering Regulation (AMLR),⁶⁶ CDD processes are a set of measures that financial institutions must implement to identify, verify, and monitor the identity and activities of their customers to assess risks related to money laundering and terrorist financing. CDD measures require financial institutions to identify and verify customer identities using various data sources, including personal data;⁶⁷ this applies to both new and ongoing business relationships.⁶⁸

⁶⁰Baden-Württemberg Commissioner for Data Protection and Freedom of Information, 'Legal Bases in Data Protection for AI' *Baden-Württemberg Data Protection* (7 November 2023) <https://www.baden-wuerttemberg.datenschutz.de/legal-bases-in-data-protection-for-ai/#_ftnref17> accessed 30 October 2024.

⁶¹This aligns with EDPB's reasoning outlined in EDPB (n 53) paras 120–23.

⁶²FATF, *International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation* (Paris, 2023) 19.

⁶³Astrid Bertrand others, 'Do AI-based Anti-money Laundering (AML) Systems Violate European Fundamental Rights?' (2021) 11 *International Data Privacy Law* 288.

⁶⁴Fenergo, 'Global Financial Institution AML and Regulatory Fines Soar in 2023' *Fenergo* (10 January 2024) <<https://resources.fenergo.com/newsroom/global-financial-institution-aml-and-regulatory-fines-soar-in-2023>> accessed 27 February 2025.

⁶⁵FATF (n 15) 19.

⁶⁶Regulation (EU) 2024/1624 of the European Parliament and of the Council of 31 May 2024 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing [2024] OJ L 1624.

⁶⁷FATF (n 62) 14.

⁶⁸*Ibid.*

Traditional AML systems often rely on static models that become outdated as fraud patterns evolve.⁶⁹ As criminal methods grow more sophisticated, the ability of legacy systems to adapt diminishes, making it increasingly difficult to distinguish between legitimate transactions and money laundering activities.⁷⁰ In light of this situation, financial institutions – known for their readiness to embrace the technological trends⁷¹ – coupled with the Financial Action Task Force (FATF) emphasis on a risk-based approach, as the foundation of effective AML systems,⁷² have been driven to adopt innovative solutions to enhance compliance. Traditionally, financial organisations were required to manually collect and verify customer information by gathering and analysing data from multiple sources, reviewing past alerts, and conducting public domain searches for adverse information on customers and their associates.⁷³ The introduction of automation and AI has streamlined these processes, resulting in faster verification and considerable time savings by reducing manual document handling.⁷⁴

Identifying suspicious financial transactions often requires detailed information about customers, their behaviours, and transaction histories. This has led banks to make substantial investments in advanced technologies, such as AI systems, due to their outstanding ability to fulfil these tasks by enhancing operational efficiency, improving predictive accuracy, lowering operational costs,⁷⁵ and helping institutions meet AML compliance requirements and avoid severe penalties associated with non-compliance.⁷⁶

Money laundering typically occurs in three stages: placement, layering, and integration.⁷⁷ In the placement stage, illegal funds get channelled into a financial institution, often through deposits or transfers.⁷⁸ During the layering stage, the launderer will attempt to move these funds across different entities or accounts using methods like wire transfers or money orders in an attempt to make the illegal funds trail harder to follow.⁷⁹ In the final stage, integration, the laundered money is used to acquire legitimate assets or fund businesses, blending with the legal economy.⁸⁰ AI could be applied to identify suspicious transactions during placement and layering, as financial institutions actively monitor these stages. The integration stage, however, is more challenging to detect, as the funds have already bypassed many detection measures.⁸¹

AML tools adopted by financial institutions follow a structured process that connects data sources directly to the monitoring system. Depending on existing data, professionals responsible for developing AI-based AML systems can establish parameters that identify

⁶⁹FATF (n 15) 11.

⁷⁰Goldbarsht (n 24) 55.

⁷¹Pedro Maia, 'Intelligent Compliance' in Maria João Antunes and Susana Aires de Sousa (eds), *Artificial Intelligence in the Economic Sector: Prevention and Responsibility* (Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2021) 16.

⁷²FATF (n 15) 13.

⁷³P Ryan and others, 'GDPR Compliance Tools: Best Practice from RegTech' in Joaquim Filipe and others (eds), *Enterprise Information Systems. ICEIS 2020* (Lecture Notes in Business Information Processing, vol 417, Springer, 2021) <https://doi.org/10.1007/978-3-030-75418-1_41> accessed 7 May 2025 918.

⁷⁴Ibid.

⁷⁵Goldbarsht (n 24) 55.

⁷⁶Ibid.

⁷⁷Kevin Sullivan, *Anti-Money Laundering in a Nutshell* (Apress, US, 2015) 7.

⁷⁸Ibid.

⁷⁹Ibid 8.

⁸⁰Jingguang Han and others, 'Artificial Intelligence for Anti-Money Laundering: A Review and Extension' (2020) 2 *Digital Finance* 216.

⁸¹Ibid.

potentially fraudulent activities,⁸² allowing the system to flag suspicious transactions for further review by human analysts.⁸³ During assessing the legitimacy of flagged transactions, analysts review customers' personal data in detail, including transaction history, account numbers, transaction information, identities and background,⁸⁴ showcasing the importance of safeguarding this information against unauthorised access, misuse, and exploitation.⁸⁵

Equally importantly, an AML system typically comprises four key layers: the data layer, the screening and monitoring layer, the alert and event layer, and the operational layer.⁸⁶ During the data layer, which serves as the foundation for identifying potential suspicious activities by collecting, managing, and storing accurate customer data, financial institutions process both internal and external data.⁸⁷ At this stage, AI systems can rely on the processed data to interpret and analyse human behaviour to gain deeper insights into the true intent behind customer transactions.⁸⁸

The second layer, screening and monitoring, automates the analysis of clients and transactions in search of potential suspicious behaviour. This stage relies extensively on the data processed in the initial layer. At this stage, AI can be used to identify correlations and anomalies indicative of money laundering activities.⁸⁹ However, with extensive monitoring comes the risk of unauthorised access or misuse of sensitive personal information. Companies providing AI systems to industry partners may be tempted to build hidden pathways, also known as 'backdoors', that enable unauthorised individuals to access the AML systems.⁹⁰ Such actions can be completed without adhering to standard security protocols to report suspicious activity, undermining the integrity of the AML systems and compromising data protection principles.

When a suspicious activity is flagged, it is escalated to the alert and event layer for human review.⁹¹ At this stage, a human operator evaluates the flagged activity and, as part of the final layer called the operational layer, determines whether to approve, reject, or block the transaction manually based on insights derived from previous layers.⁹² Summarising, by incorporating AI, AML systems can better build detailed profiles of customers, predict potential risks, and monitor deviations from standard financial behaviours. This enables banks to intercept these illegal activities as they occur, while also forecasting potential suspicious behaviour in the future.

⁸²Rashid Alhajeri and Abdulrahman Alhashem, 'Using Artificial Intelligence to Combat Money Laundering' (2023) 16 *Intelligent Information Management* 293.

⁸³Jingguang Han and others, 'Artificial Intelligence for Anti-money Laundering: A Review and Extension' (2020) 2 *Digital Finance* (Springer) 216.

⁸⁴Ezekiel Onyekachukwu Udeh and others, 'The Role of Big Data in Detecting and Preventing Financial Fraud in Digital Transactions' (2024) 22 *World Journal of Advanced Research and Reviews* 1753.

⁸⁵*Ibid.*

⁸⁶Alhajeri (n 82) 293.

⁸⁷Jingguang Han and others, 'Artificial Intelligence for Anti-money Laundering: A Review and Extension' (2020) 2 *Digital Finance* (Springer) 216.

⁸⁸Rashid Alhajeri and Abdulrahman Alhashem, 'Using Artificial Intelligence to Combat Money Laundering' (2023) 16 *Intelligent Information Management* 294.

⁸⁹Han and others (n 87) 219.

⁹⁰The Commodity Futures Trading Commission, '*Responsible Artificial Intelligence in Financial Markets*' (Opportunities, Risks and Recommendations, Report of the subcommittee on Emerging and Evolving Technologies, 2024) 44.

⁹¹Han and others (n 87) 219.

⁹²*Ibid.* 220.

3.2. Functions of AI in AML/CFT

AI systems can be used for several purposes in the AML/CFT arena. For the purposes of this article, four procedures will be specifically analysed: Customer Due Diligence, Risk Profiling and Scoring, Transaction Monitoring, and Suspicious Activity Reporting.

3.2.1. Customer due diligence (CDD)

Due Diligence is generally defined as *'the processes through which enterprises can identify, prevent, mitigate and account for how they address their actual and potential adverse impacts'*.⁹³ Although there is no uniform definition in the European AML/CFT rules and regulations, the first time that the notion of 'Customer Due Diligence' appeared was in 2005, through the 3rd Anti-Money Laundering Directive.⁹⁴

In the 2024 AML package, Customer due diligence measures are determined in Article 20 of the AMLR. More specifically, according to Article 20(1) AMLR, to comply with their CDD obligations, all obliged entities must proceed to a series of measures, such as – inductively – the identification of a customer and the verification of their identity, the assessment and information retrieval about the customer's beneficial ownership or on the nature of the customer's business, as well as the identity of third parties conducting transactions on behalf of a customer. The AMLR does not prohibit the use of AI systems to carry out these procedures. Instead, Article 20(2) AMLR stipulates that

obliged entities shall determine the extent of the measures ... on the basis of an individual analysis of the risks of money laundering and terrorist financing having regard to the specific characteristics of the client and of the business relationship or occasional transaction and taking into account the business-wide risk assessment by the obliged entity.

Therefore, under no means is the use of AI systems prohibited in these procedures.

Integral to the broader obligation of conducting CDD measures, the most crucial procedure is conducting detailed background checks to gather sufficient information about customers and assess their potential involvement in illicit activities.⁹⁵ Financial institutions are required to verify identification details provided by customers at the start of any financial relationship. Consequently, they may refrain from establishing a business relationship if a client fails to verify their identity or if their background raises suspicions.⁹⁶ For that reason, processing personal data during CDD is essential to evaluate the risks posed by customers, not only to the institution itself but also to society in general.⁹⁷

The enhancement of CDD protocols is one of the most critical applications of AI in combating money laundering and improving customer identification. AI systems significantly enhance identity verification processes by analysing information from various sources.⁹⁸ Traditionally, CDD procedures required the manual verification of customer documents, which is undeniably time-consuming and demands significant resources in terms of

⁹³OECD Guidelines for Multinational Enterprises, Chapter II –General Policies, 10.

⁹⁴Directive 2005/60/EC of the European Parliament and of the Council of 26 October 2005 on the prevention of the use of the financial system for the purpose of money laundering and terrorist financing

⁹⁵Dennis Cox, *Handbook of Anti-money Laundering* (John Wiley & Sons, 2014) 152.

⁹⁶Ibid 153.

⁹⁷Ibid 169.

⁹⁸Abhishek Gupta and others, 'Applying Machine Learning for Effective Customer Risk Assessment' (2023) *Artificial Intelligence Applications in Banking and Financial Services*. Future of Business and Finance. Springer Singapore, 66–67.

time and labour.⁹⁹ The integration of AI has helped streamline this process, increasing both efficiency and accuracy. Subsequently, the data workflow transitions to customer risk-profiling, wherein AI systems leverage collected personal data to evaluate and classify customers based on their risk levels.

3.2.2. Customer risk profiling and scoring

Risk Scoring is provided in the new AMLR, more specifically in the context of enhanced CDD.¹⁰⁰ Once the data is collected, AI-based AML systems may generate risk profiles for each customer by analysing transaction history, geolocation, and behaviour patterns. Personal data such as income levels, transaction frequency, and known associations with high-risk entities or regions play a significant role in determining risk levels.¹⁰¹ Based on these factors, AI systems can classify clients into low-, medium-, or high-risk categories. In segmenting customers based on risk profiles,¹⁰² AI enables financial institutions to take preventive measures in real-time, effectively addressing potential suspicious activities before they escalate.¹⁰³ This risk profiling paves the way for the next critical function: transaction monitoring and anomaly detection, where AI continuously assesses transaction activities for suspicious behaviour.

3.2.3. Transaction monitoring and anomaly detection

Machine learning algorithms can be leveraged by AI-driven AML systems to monitor conducted transactions effectively. As they process large amounts of data originating from various sources – such as customer accounts and past financial transactions¹⁰⁴ – they can identify anomalies and analyse customer behaviour patterns, such as transaction habits and timings, that could indicate money laundering.¹⁰⁵ AI algorithms can also be employed to analyse personal documents, such as passports or driving licences, to better verify identities and more effectively detect suspicious transactions.¹⁰⁶ Traditional AML approaches, which relied on manual reviews, were limited in their capacity to manage the increasing complexity and volume of financial transactions, unlike AI-based systems, which can almost instantly process vast datasets to uncover complex patterns of suspicious activity while adapting to new money laundering schemes.¹⁰⁷ For example, geolocation data can help identify suspicious activities by revealing the physical locations of transactions. Hence, personal data, when combined with other data points – such as device information – used by AI-based AML systems, may identify abnormal activities or patterns that may suggest money laundering activity.¹⁰⁸

⁹⁹Paul Ryan and others, 'GDPR Compliance Tools: Best Practice from RegTech' (2021) *Enterprise Information Systems*, 918.

¹⁰⁰AMLE Article 20(2).

¹⁰¹Mahiya Raj and others, 'The Use of Artificial Intelligence in Anti-money Laundering (AML)' (2024), 2024 3rd International Conference on Sentiment Analysis and Deep Learning 274–75.

¹⁰²Magdalena Brewczyńska and Eleni Kosta, 'From the Fight against Money Laundering and Financing of Terrorism towards the Fight for Fundamental Rights: Role of Data Protection' (2023) 11.

¹⁰³Oracle, 'Fight Money Laundering More Effectively with AI-Enhanced Tools' <<https://www.oracle.com/nl/financial-services/aml-ai/>> accessed 15 March 2025.

¹⁰⁴Faheem Ashraf and Alejandro Schaffer, 'Combating Financial Crime: AI and Machine Learning in Anomaly Detection and Risk Management' (2024) <<https://rgdoi.net/10.13140/RG.2.2.22450.41927>> accessed 15 March 2025.

¹⁰⁵Raj and others (n 101) 274

¹⁰⁶Ibid.

¹⁰⁷Yasir Nawaz and Sebastian Thrun, 'Integrating AI and Data Mining for Smarter Anti-money Laundering (AML) Solutions' (2024) <<https://doi.org/10.13140/RG.2.2.18675.54569>>.

¹⁰⁸FATF, *Illicit Financial Flows from Cyber-enabled Fraud* (2023) 68

Using personal data for real-time transaction monitoring is an important feature of AI-based AML systems. Financial transactions include a variety of information, which covers the amount, frequency, destination, and purpose of transfers.¹⁰⁹ For example, if a customer who usually performs small domestic transactions abruptly starts a large international transfer to a high-risk jurisdictions – countries known for having significant gaps in their AML/CFT financing efforts¹¹⁰ – the AI-based system may flag the activity as suspicious.¹¹¹ Using pattern recognition techniques, AI-powered systems assess in real time the transaction amounts, destinations, and timestamps relative to the customer's historical activity, helping to detect potential money laundering attempts.¹¹² AI systems can also be trained on past transaction data to identify patterns that resemble known money laundering techniques¹¹³ allowing flagged transactions to undergo further investigation.¹¹⁴

3.2.4. Suspicious activity reporting

Spotted deviations lead to the generation of Suspicious Transaction Reports (STRs), which obliged entities are required to submit to competent authorities (such as Financial Intelligence Units) upon identifying potential money laundering attempts.¹¹⁵ STRs typically contain personal data, such as customer names, transaction details, and related risk factors.¹¹⁶ Employment of AI-powered systems can automate much of this process by generating alerts when suspicious activities are identified and automatically creating and submitting STRs to competent authorities, speeding up the reporting process.¹¹⁷ This not only accelerates the reporting process but also reduces the burden on compliance teams.¹¹⁸

4. Applicability of legal grounds under Article 6(1) of the GDPR for processing personal data to train AML AI systems

This paper primarily aims to verify whether the legal basis of legitimate interest [Article 6(1)(f) GDPR] can be utilised to train AI-enhanced AML/CFT tools on personal data. Any processing of personal data must meet at least one of the six bases in Article 6(1) of the GDPR. Before addressing the legal basis of legitimate interest, it should be determined why the other grounds in Article 6(1) of the GDPR would not be feasible.

¹⁰⁹International Monetary Fund, 'International Transactions in Remittances: Guide for Compilers and Users' (2009) 28.

¹¹⁰FATF, 'High-risk Jurisdictions subject to a Call for Action – 25 October 2024' <<https://www.fatf-gafi.org/en/publications/High-risk-and-other-monitored-jurisdictions/Call-for-action-october-2024.html>> accessed 28 November 2024.

¹¹¹Raj and others (n 101) 274.

¹¹²Financial crime academy, 'Enhancing Compliance: The Advantages of AI-based Transaction Monitoring' *Financial crime academy* (22 July 2024), <<https://financialcrimeacademy.org/ai-based-transaction-monitoring/>> accessed 28 August 2024.

¹¹³Brewczyńska and Kosta (n 102).

¹¹⁴Raj and others (n 101) 274.

¹¹⁵FATF *International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation, Recommendation* (2012–2025) 20.

¹¹⁶AMLR, Article 69(2).

¹¹⁷Kennedy Meda, 'Revolutionizing SARs Filings and Investigation Efficiency through AI' *Thomson Reuters* (2024) <<https://www.thomsonreuters.com/en-us/posts/investigation-fraud-and-risk/ai-revolution-sars-filings/>>

¹¹⁸*ibid.*

While consent [Article 6(1)(a) GDPR] may seem like a potentially viable option at first glance,¹¹⁹ it is impractical for AML purposes. Obtaining it from every concerned individual beforehand would be burdensome. Additionally, consent can be withdrawn at any time, making it an unstable foundation for an ongoing task driven by public interest, such as AML, and posing technical challenges, particularly due to the difficulties in erasing data once entered into the training stage.¹²⁰

The performance of a contract legal basis [Article 6(1)(b) GDPR] applies only when processing is necessary to fulfil contractual obligations ‘or in order to take steps at the request of the data subject before entering into a contract’. The EDPB position is that to apply the contractual basis, there must be a ‘necessity’ between the service and the processing.¹²¹ This reasoning is reinforced by the CJEU in the *Meta* case (C-252/21), ruling that

the linking (...) can be regarded as necessary for the performance of a contract (...) only on condition that the processing is objectively indispensable for a purpose that is integral to the contractual obligation intended for those users, such that the main subject matter of the contract cannot be achieved if that processing does not occur.¹²²

Against this background, processing personal data for the purposes of AI-enhanced AML tools cannot be considered indispensable for fulfilling contracts with customers. Instead, it assists obliged entities in complying with legal obligations under AML legislation. Although such processing can be claimed to be functionally associated with the performance of contracts, it is too remote to meet the high standard of linkage established by the CJEU. Therefore, this legal basis could not be feasibly used as a basis for such processing.

Processing based on vital interests [Article 6(1)(d) GDPR] is generally reserved for situations posing an imminent risk to someone’s life or safety. It is interpreted narrowly as an exception and refers to cases of life-threatening and severe health risks (‘essential for the life’).¹²³ This is irrelevant in this context, as AML efforts focus on preventing financial abuses rather than ensuring immediate personal safety.

¹¹⁹On the notion and conditions for valid consent: EDPB, ‘Guidelines 05/2020 on consent under Regulation 2016/679, version 1.1 (4 May 2020)’ <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf> accessed 27 December 2024.

¹²⁰Cf. the analysis in the section 2.

¹²¹EDPB, ‘Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subjects’ (8 October 2019) <https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf> accessed 27 December 2024 para 50.

¹²²Case C-252/21 *Meta v Bundeskartellamt* [2023] ECLI:EU:C:2023:537, para 125.

¹²³Recital 46 GDPR:

The processing of personal data should also be regarded to be lawful where it is necessary to protect an interest which is essential for the life of the data subject or that of another natural person. Processing of personal data based on the vital interest of another natural person should in principle take place only where the processing cannot be manifestly based on another legal basis. Some types of processing may serve both important grounds of public interest and the vital interests of the data subject as for instance when processing is necessary for humanitarian purposes, including for monitoring epidemics and their spread or in situations of humanitarian emergencies, in particular in situations of natural and man-made disasters.

The legal basis for performing public tasks [Article 6(1)(e) GDPR]¹²⁴ applies when processing is necessary for tasks carried out in the public interest or the exercise of official authority, typically by public bodies.¹²⁵ This basis would be at least theoretically worth considering following CNIL's position on AI,¹²⁶ provided that two conditions are met. First, a given task should be stipulated in the statutory law applicable to the controller. Secondly, training AI-enhanced AML tools should make it 'possible to carry out this task specifically (...) in a relevant and appropriate manner'. CNIL excludes situations where such processing 'pursues an objective (...) unrelated to it or too far removed from its particularities'. There are at least two arguments against using this basis. Firstly, banks and financial institutions are primarily private entities that do not perform public tasks, so they cannot rely on this basis by their nature. Even if this basis could be considered by these entities performing public tasks, they would be relatively rare as it requires specific provisions established in legal acts following Article 6(3) of the GDPR requirements. In such circumstances, depending on the wording of such provision, it would be debatable whether (1) processing personal data for training AI-enhanced AML tools is necessary and not too remote from the underlying public task; (2) such provision is proportionate to the legitimate aim pursued in the first place since less restrictive (i.e. without AI-enhanced tools) actions might achieve the same objectives.

Notwithstanding the above, the legal basis of Article 6(1)(c) GDPR is of relevance in this discourse. According to Article 6(1)(c) GDPR, processing is lawful if it is necessary for compliance with a legal obligation to which the controller is subject. Processing personal data on a 'legal obligation' basis is lawful only if is necessary to fulfil that obligation. Financial institutions are required to handle customer data to comply with regulatory requirements, particularly for CDD purposes. This includes various processing activities, such as collecting, storing, assessing, and, when necessary, sharing data to verify identities and evaluate risks. However, while they are legally obliged to process such data, they are not required to use AI systems to achieve this goal; the use of AI remains entirely optional. Although it seems plausible that national or EU law may introduce explicit legal obligations to use AI for various purposes at some point,¹²⁷ this is not the case at present. Reliance on this basis would require the obligation to use AI to be explicitly indicated in the legal act's text or refer to AI functionality in such a way that the obligation would be unfulfillable otherwise. Currently, AI systems are employed by obliged entities for efficiency purposes, mainly to save resources, workforce, and time, but it does not constitute a strict legal necessity. Therefore, relying on this legal basis is questionable, as obliged entities may have difficulties establishing a clear link of necessity for fulfilling legal obligations.¹²⁸

¹²⁴Read in conjunction with Article 6(3) GDPR. Cf. also Council of Europe, *Guidelines on the Risk-Based Approach to Combating Money Laundering and Financing of Terrorism (AML/CFT) for Supervised Entities in the Legal Professions* (Council of Europe, August 2023) 18–20 <<https://rm.coe.int/prems-109823-gbr-2051-guidelines-aml-cft-a5-web/1680b156d9>> accessed 31 December 2024.

¹²⁵CNIL, *Ensuring the Lawfulness of the Data Processing* (2024) <<https://www.cnil.fr/en/ensuring-lawfulness-data-processing>> accessed 27 December 2024.

¹²⁶Ibid.

¹²⁷Although meeting standards of establishing such legal obligation following Article 6(3) of the GDPR requirements might be challenging, cf. Eyup Kun, 'Searching for the Appropriate Legal Basis for Personal Data Processing for Cybersecurity Purposes under the NIS 2 Directive: Legal Obligation and/or Legitimate Interest?' (2025) 56 *Computer Law & Security Review* 106098 8–9.

¹²⁸Similar arguments in the context of cybersecurity law in: Ibid 8–9.

5. Legitimate interest [Article 6(1)(f) GDPR] as a legal ground for AML purposes

This section of the paper verifies whether the legal basis of legitimate interest [Article 6 (1)(f) of the GDPR]¹²⁹ could serve as an alternative legal basis for training AI-enhanced AML tools. This legal basis is highly evaluative.¹³⁰ It requires satisfying a complex three-stage test: (1) identifying a legitimate interest of the controller or a third party; (2) demonstrating the necessity of processing; (3) balancing it against the individual's interests, rights, and freedoms.¹³¹ The legal basis for processing, including legitimate interest, has recently been the subject of increased interest by the DPAs, who have issued a number of papers and opinions,¹³² culminating in EDPB 'Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models',¹³³ the analysis of which is vital to this section.¹³⁴ This EDPB opinion is characterised by a high level of generality and frequently employs hedging terms (e.g. may), leaving ample room for national DPAs to assess situations on a case-by-case basis.¹³⁵

¹²⁹Article 6(1)(f) GDPR:

Processing shall be lawful only if (...) is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except 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.

Cf. Recital 47 GDPR.

¹³⁰R Niedermeier and ME Mpame, 'Processing Personal Data under Article 6 (f) of the GDPR: The Concept of Legitimate Interest' [2019] 3 International Journal for the Data Protection Officer, Privacy Officer and Privacy Counsel 18.

¹³¹Article 29 Working Party, 'Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC (09.04.2014)' <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf> accessed 27 December 2024; EDPB, 'Guidelines 1/2024 on Processing of Personal Data Based on Article 6(1)(f) GDPR. Version 1.0' (8 October 2024) <https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf> accessed 3 January 2025 para 6; The Information Commissioner's Office, 'Legitimate interests' *The Information Commissioner's Office* (19 May 2023) <<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/legitimate-interests/>> accessed 27 December 2024.

¹³²EDPB, 'Report of the Work Undertaken by the ChatGPT Taskforce' (23 May 2024) <https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf> accessed 3 January 2025; EDPB, 'Generative AI and the EUDPR: First EDPS Orientations for Ensuring Data Protection Compliance When Using Generative AI Systems' (3 June 2024) <https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf> accessed 3 January 2025; Data Protection Authority of Baden-Württemberg, 'Rechtsgrundlagen im Datenschutz beim Einsatz von Künstlicher Intelligenz' (17 October 2024) <<https://www.baden-wuerttemberg.datenschutz.de/rechtsgrundlagen-datenschutz-ki/>> accessed 3 January 2025; The Hamburg Commissioner for Data Protection and Freedom of Information, 'Discussion Paper: Large Language Models and Personal Data' (2024) <https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/240715_Discussion_Paper_Hamburg_DPA_KI_Models.pdf> accessed 3 January 2025; Data Protection Authority of Belgium, 'Artificial Intelligence Systems and the GDPR: A Data Protection Perspective' (September 2024) <<https://www.autoriteprotectiondonnees.be/publications/artificial-intelligence-systems-and-the-gdpr---a-data-protection-perspective.pdf>> accessed 3 January 2025; Data Protection Commission, Ireland, 'AI, Large Language Models and Data Protection' (18 July 2024) <<https://www.dataprotection.ie/en/dpc-guidance/blogs/AI-LLMs-and-Data-Protection>> accessed 3 January 2025; CNIL, 'AI System Development: CNIL's Recommendations to Comply with the GDPR' (7 June 2024) <<https://www.cnil.fr/en/ai-system-development-cnils-recommendations-comply-gdpr>> accessed 3 January 2025. Commission Nationale de l'Informatique et des Libertés (CNIL), France, 'Ensuring Lawfulness of Data Processing: Legal Basis' (2024) <<https://www.cnil.fr/en/ensuring-lawfulness-data-processing-legal-basis>> accessed 3 January 2025.

¹³³EDPB (n 53).

¹³⁴This opinion addresses not only the issue of the applicability of legitimate interest as a legal basis in the AI lifecycle but also the question of whether an AI model can be 'anonymous and the consequences of unlawful personal data processing during the development phase on the subsequent operation of the AI model.

¹³⁵EDPB Opinion refers to AI models, which are distinguishable from AI systems, the definitions of which were presented in Section 2 of this paper. We do not analyse the relationship of AI models to AI systems, taking as EDPB Opinion that an

AI system will rely on an AI model to perform its intended objective by incorporating the model into a larger framework (e.g. an AI system for customer service might use an AI model trained on historical conversation data to provide responses to user queries). (para 22)

The first stage is to identify the legitimate interest of the controller or a third party. In the literature in the context of text mining, it is highlighted that such a legitimate interest of the controller could potentially be derived from Article 16 CFR¹³⁶ concerning the freedom to conduct business.¹³⁷ Different interests deserve different levels of protection – those that are more compelling (i.e. generally serving society), and those less so (i.e. generally serving more individual interests).¹³⁸ In its recent judgment, the CJEU confirmed that commercial interest as such can also constitute a legitimate interest of the controller.¹³⁹ To meet the requirements of being legitimate, the interest must be lawful, clearly expressed, and not speculative.¹⁴⁰ This holds for AML/CFT purposes; the AMLR (in Recital 150) explicitly acknowledges AML/CFT as a significant public interest. Therefore, obliged entities have a clear, real, and lawful interest in this scenario. By utilising automated tools like AI, they can swiftly identify suspicious patterns, saving time and labour. From society's perspective, this is also advantageous as speed is crucial in detecting money laundering and terrorist financing, reducing the opportunities for criminals to benefit from their illicit activities and hindering terrorist organisations.

The next step is demonstrating that processing is necessary to achieve the identified legitimate interest. The assessment criterion is whether there is a reasonable and effective but 'less intrusive' way to achieve the legitimated interest,¹⁴¹ for example, by using anonymised data. It can be argued that the complete functionalities of AI-enhanced AML tools can only be achieved by including personal data in training datasets. If obliged entities decide to use AI-enhanced AML tools, high-quality data¹⁴² (including personal data) is needed at the training stage to increase the correctness of the results of such tools and reduce their bias, which poses a serious risk in the AML/CFT context.¹⁴³ However, this places the onus on those who train AI tools to make a trade-off between achieving these goals and precisely determining the extent of the data required to achieve them.¹⁴⁴

The critical challenge lies in the third stage, namely the balancing act of the legitimate interest of the controller (or third party) against the individual's interests, rights, and freedoms.¹⁴⁵ Like the controller's legitimate interests, individuals' interests, rights, and freedoms can vary considerably.¹⁴⁶ Factors to be taken into account in this test are: (1)

¹³⁶Article 16 of the Charter of Fundamental Rights: 'The freedom to conduct a business in accordance with Community law and national laws and practices is recognised'.

¹³⁷Benjamin Bremert, 'Legal Aspects of Text Mining Publicly Available Data' *Unabhängiges Landeszentrum für Datenschutz* (7 September 2017) 5–6, accessed 27 December 2024.

¹³⁸Article 29 Working Party (n 131) 24; EDPB (n 131) 17.

¹³⁹Case C-621/22 *Koninklijke Nederlandse Lawn Tennisbond v Autoriteit Persoonsgegevens* [2024] ECLI:EU:C:2024:857.

¹⁴⁰Article 29 Working Party (n 131) 25; EDPB (n 131) 17.

¹⁴¹C Altobelli, 'To Scrape or Not to Scrape? The Lawfulness of Social Media Crawling under the GDPR' in J Herveg (ed), *Deep Diving into Data Protection: 1979–2019 Celebrating 40 Years of Privacy and Data Protection at the CRIDS* (Éditions Larcier 2021) 12, 151–75 <https://www.researchgate.net/publication/351227024_To_Scrape_or_Not_to_Scrape_The_Lawfulness_of_Social_Media_Crawling_under_the_GDPR> accessed 27 December 2024.

¹⁴²Goldbarsht (n 24) 62–63.

¹⁴³Goldbarsht (n 24) 64–65; Georgios Pavlidis, 'Deploying Artificial Intelligence for Anti-money Laundering and Asset Recovery: The Dawn of a New Era' (2023) 26(7) *Journal of Money Laundering Control* 159.

¹⁴⁴EDPB (n 131) 29–30.

¹⁴⁵J Kamara and P De Hert, 'Understanding the Balancing Act behind the Legitimate Interest of the Controller Ground: A Pragmatic Approach' [2018] 4(12) *Brussels Privacy Hub* <<https://ssrn.com/abstract=3228369>> accessed 27 December 2024.

¹⁴⁶Article 29 Working Party (n 131) 30–31; EDPB (n 131) 32, 35–38.

data subjects' interests, fundamental rights, and freedoms; (2) the impact of processing, considering the nature and the context of personal data processing and further consequences, including likelihood and depth of the risk¹⁴⁷; and (3) the reasonable expectations of the data subjects.¹⁴⁸ As indicated above, the legitimate interest in enhancing AML/CFT protection is recognised in law as a public interest. Therefore, a key concern is balancing this with the position of data subjects. We analyse each criterion below in the context of AI-enhanced AML tools. It should be noted that these criteria are functionally linked to each other and inextricably interact.

Data subjects' interests, fundamental rights and freedoms are broad concepts and vary depending on the context. They provide a benchmark for assessing the chances and risks of given personal data processing. They also include retaining control over their personal data by data subjects.¹⁴⁹ Within the context of training for AML tools, data subjects may have an interest in retaining control over their personal data, which can vary based on whether data is sourced from first parties or third parties and whether they are later subject to decisions made with the support of these tools. The boosted capacity of obligated entities to infer data using AI-enhanced tools exposes individuals to greater intrusiveness by facilitating the identification of previously unseen correlations. Being incorrectly flagged by an AI-enhanced tool could result in frozen accounts or delayed transactions. This issue is further complicated by the 'black box' phenomenon, potentially leaving individuals without clear answers regarding inference methods. On the other hand, it may benefit data subjects as financial services customers by minimising money laundering and terrorist financing, improving overall economic safety.

The impact of the processing on the data subjects – linked with their identified interests in the AML/CFT scenario – is assessed, as indicated above, given the nature and context of personal data processing and its further consequences. Regarding the nature of the data,¹⁵⁰ it should be noted that it is not feasible to categorise all personal data into a single group. Even data not classified as specific categories [under Article 9(1) of the GDPR] may vary in sensitivity. In the context of AML, the focus is primarily on financial data, which can disclose information about an individual that may significantly impact their private and professional life, including overseas trips they have taken, spending habits, or even, indirectly, addictions (e.g. regular purchases at the liquor shop). Not all data for these tools is obtained directly from the data subjects; some may also be sourced from publicly available materials such as public registers or news reports, which should be evaluated in light of the potential for unexpected connections inferred by such tools.

In the context of using personal data to train AI-driven AML tools,¹⁵¹ it should be emphasised that the evaluation depends on how these tools are designed, considering the scopes of data, safeguards, and their intended use. For example, employing these tools for continuous monitoring is more intrusive than a one-off check. Undoubtedly, it is helpful that AI training occurs mainly within financial institutions (or their third-party providers) that are already subject to regulatory scrutiny regarding data protection and cybersecurity, which enhances the likelihood of appropriate data handling. However,

¹⁴⁷Article 29 Working Party (n 131) 33–41; EDPB (n 131) 32, 45–49.

¹⁴⁸Article 29 Working Party (n 131) 36; EDPB (n 131) 32, 39.

¹⁴⁹EDPB (n 53) 77–81.

¹⁵⁰Ibid 84.

¹⁵¹Ibid 85–86.

this does not eliminate challenges. The difficulties surrounding the transparency of AI-enhanced tools could have consequences for data subjects and their reasonable expectations. It can also be argued that it would be difficult (due to the probes with the erasure of data)¹⁵² to introduce an effective opt-out mechanism, which impacts the ability of data subjects to exercise their right to object.¹⁵³

Further consequences encompass the type, depth, and likelihood of risks for data subjects.¹⁵⁴ The processing in one phase of the AI cycle can lead to consequences that emerge in other phases. The impacts revealed during the use of AI-enhanced tools by obliged entities may relate to the training or re-training of those tools. It requires consideration of what types of risks, apart from the loss of control over one's data, should be taken into account. Firstly, false positives – the incorrect identification of individuals or transactions as suspicious – can result in financial exclusion (as indicated above). In such scenarios, individuals may suffer reputational harm when incorrectly linked to criminal activity. There may also be potential discriminatory effects if AI tools display biased patterns against specific groups. Moreover, the technical opacity of AI makes error correction challenging, as incorrect data or decisions become embedded. The likelihood of these risks depends on the organisational and technical measures put in place as safeguards, which will be discussed later in this section. It can be argued that with these measures, there is an opportunity to minimise these risks, which could also materialise without the implementation of AI.

The EDPB identifies the assessment of reasonable expectations as crucial in the balancing test because of the complex and obscure nature of AI as technology and the potential various use cases, underlying that mere information about training in the controller's privacy policy might not be sufficient.¹⁵⁵ In the AML context, several factors should be taken into account. First, customers should generally expect some level of AML/CFT monitoring. From a mere general awareness of these operations, it is not yet possible to derive a reasonable expectation that the individuals concerned will be used to train such tools. Even so, the extent of AI-enhanced tools' capabilities may exceed what customers reasonably expect. Second, the assessment of reasonable expectation depends on data sources (first-party or third-party) and the proximity of the purpose of initial processing to AML/CFT operations. For example, personal data from public records is likely to be associated with AML/CFT operations with less predictability, whereas if the data subject is a person opening a bank account, then the expectations of such processing may be deemed higher. Expectations may differ based on the date of data acquisition (before or after the proliferation of AI tools), customer type (individuals and businesses), transaction types (domestic and international), and prior experience with AML procedures within an obliged entity. Finally, the opacity of AI systems makes it difficult for customers to understand how their data is processed, rendering the assessment of reasonable expectations more challenging for obliged entities.

The above analysis shows that each criterion is highly evaluative and cannot be clearly stated in favour of or against the balancing. It seems reasonable to argue that introducing extensive safeguards, such as privacy-enhancing techniques, influences the likelihood of

¹⁵²As discussed in Section 2.

¹⁵³Article 21 GDPR.

¹⁵⁴EDPB (n 53) 87–90.

¹⁵⁵*Ibid* 91–95.

passing the test.¹⁵⁶ Such safeguards are debated (extensively, albeit not exhaustively) in the EDPB Opinion, which differentiates between safeguards in the development and deployment phases.¹⁵⁷ In training AI-enhanced AML/CFT tools, we focus on safeguards during development. Training AI models for such tools should include analysing the sources and scope of the obtained data, followed by data preparation techniques, including anonymisation, pseudonymisation, and data minimisation strategies, like masking personal data; and the training process should employ privacy-preserving techniques such as differential privacy.¹⁵⁸

In addition, further procedural safeguards merit consideration. The first step is to implement a waiting period between data collection and AI training, allowing data subjects to exercise their rights before their data is utilised for training. The second step involves an opt-out right surpassing GDPR requirements, providing data subjects with greater control over their data before AI training. Furthermore, expanded erasure rights could introduce additional flexibility by widening the circumstances under which erasure is possible. Moreover, the procedures for identifying and mitigating false positives through enhanced human review and ‘unlearning’ strategies via retraining are worth implementing. Additionally, the EDPB advocates for increasing the transparency of AI-enhanced tools not just by updating privacy policies but also by launching dedicated awareness campaigns aimed at data subjects.¹⁵⁹ Ultimately, the thoughtful integration of these safeguards into the training of AI-enhanced AML tools and their timely reviews contribute significantly to meeting the balancing test favourably.

This analysis demonstrates that legitimate interest can serve as a legal basis for using personal data to train AI-enhanced AML tools; however, it necessitates a complex balancing test under Article 6(1)(f) of the GDPR. The public interest in AML/CFT measures supports this argument, but obliged entities must take into account implications for data subjects, including the sensitivity and scope of personal data used for training, the intrinsic technical limitations of AI systems, and potential difficulties in exercising data subject rights. The viability of this legal basis ultimately depends on implementing well-thought-out technical and organisational safeguards, including privacy-preserving techniques during the training phase, which also requires producing, monitoring and updating processes and documentation to demonstrate accountability.

6. Conclusion

The evolution of money laundering techniques demands advanced detection methods. However, detection must be conducted within a clear regulatory framework. While AI-enhanced tools may offer powerful tools for streamlining the detection and prevention of money laundering, the regulatory challenges leave obliged entities with doubts. This paper focuses on the issue of legal grounds for training AI-enhanced tools on personal data for AML/CFT purposes.

¹⁵⁶M Van Bekkum and F Zuiderveen Borgesius, ‘Using Sensitive Data to Prevent Discrimination by Artificial Intelligence: Does the GDPR Need a New Exception?’ [2023] 48(105770) *Computer Law & Security Review* 42.

¹⁵⁷EDPB (n 53) 96–108, 44–55. Noteworthy, safeguards affecting balancing test assessment should be distinguished from the mitigating measures that must be implemented to ensure compliance with the GDPR, EDPB (n 53) 97.

¹⁵⁸EDPB (n 53) 49–52.

¹⁵⁹EDPB (n 53) 102–03.

Our analysis concludes that the basis of the necessity of compliance with a legal obligation [Article 6(1)(c) GDPR] for training on personal data proves problematic. Processing personal data for AML/CFT purposes *per se* can indeed be based on Article 6(1)(c) GDPR. However, obliged entities training AI systems to perform this work in order to save resources should not fall under Article 6(1)(c) GDPR. No legal obligation exists to train or use AI systems for AML/CFT purposes, as the European AML/CFT regime does not explicitly mandate the use of AI-enhanced tools, rendering this basis ultimately unsuitable.

On the contrary, the legal basis of legitimate interest [Article 6(1)(f) GDPR] arises as a potential appropriate legal ground, although it requires careful consideration and a case-by-case assessment. Financial institutions can claim their interest in effective money laundering detection, while society benefits from the quicker and more efficient identification of suspicious activities. However, the necessary balancing test must carefully weigh those interests against the rights and freedoms of individuals. This may become harder taking into account the opacity of AI systems, the difficulties in establishing effective opt-out mechanisms, and the challenges in erasing data once they have been inputted at the training stage. It is worth noting that technical efforts are underway to enhance the performance of AI systems in this area.

These findings highlight the need for a mature governance framework to balance efficient AML processes, based on AI, with data protection. Such a framework requires dialogue between regulators and financial institutions to develop practical guidelines. Exploring alternative approaches for training AI-enhanced tools, such as synthetic data, is plausible.

Disclosure statement

No potential conflict of interest was reported by the author(s).