sciendo
Vietnamese Journal of Legal Sciences, Vol. 07, No. 02, 2022, pp. 39-58 DOI: 10.2478/vjls-2022-0007

ARTIFICIAL INTELLIGENCE AND DATA PROTECTION: HOW TO RECONCILE BOTH AREAS FROM THE EUROPEAN LAW PERSPECTIVE

MATTHIAS ARTZT
Senior Legal Counsel at Deutsche Bank AG Frankfurt
Email: Matthias.Artzt@db.com
TRAN VIET DUNG
Dean of the Faculty of International Law, Ho Chi Minh City University of Law
Email: tvdung@hcmulaw.edu.vn

Abstract

Artificial intelligence (AI) is the latest in a long wave of disruptive technology, offering significant benefits but also creating risks if deployed in an uncontrolled manner. Recognizing this, many countries around the world, including Vietnam, are making efforts to research and build a legal framework to regulate and manage the development of AI to ensure that this technology supports their socio-economic development. In this context, the legislative experience of the European Union, one of world's leaders in digital technology, is of high significance.

The European GDPR, adopted in 2016, has become the international standard in terms of data protection. The EU AI Act, which is currently in a draft stage and will come into force not before end of 2024, sets out horizontal rules for the development, commodification and use of AI-driven products, services and systems, including the AI system and data protection. This article particularly scrutinizes the interplay between both the AI Act/AI systems and the GDPR and considers the regulatory state of play as at October 2022.

Keywords: *Artificial Intelligence, AI Act, GDPR, European Union, data protection*

As AI increasingly features in everyday life, it's no surprise to hear calls for increased regulation of the technology. The development of legal mechanisms to address the impacts or consequences of AI is essential for all countries in the context of the explosion of the Fourth Technology Revolution. One might think that challenges on the subject matter appear facially similar to those in other technocratic domains, such as the anti-money laundering law or environmental law. But AI is unique. AI's distinctiveness comes from the specific technical attributes of speed, complexity, and unpredictability - that strained legislative approach and strategy, combined with the institutional settings and incentives that influence its development. This difference means that traditional, industry-specific approaches will not guarantee effectiveness, and will quickly reach their limits. It is necessary to have a right orientation and strategy for law-making in accordance with the development of technology.

Since implementing the Doi Moi (Renovation) reform policy in 1987, Vietnamese government has been constantly searching for new drivers of

economic growth.[1] In the age of Industrial Revolution 4.0 the government has openly acknowledged the role of AI technology and considered it as a driving force of development. In March of 2021, then Prime Minister Nguyen Xuan Phuc announced a Master Plan to develop the industry of artificial intelligence in Vietnam. The plan, entitled the "National Strategy on research, development (R&D) and Application of AI" lays out Vietnam's plan to develop AI until 2030 with the goal to turn Vietnam into the top 5 leading countries in the ASEAN and the top 60 leading countries in the world in the field of R&D and application of AI.[2]

To achieve this goal, a comprehensive legal framework for AI is required. Hence, currently, the status of the AI legal framework in Vietnam is rather modest. The lawmakers generally are quite cautious in introducing the AI regulations fearing that they may seriously alter traditional legal concepts relating to the legal liability of involved parties. Arguably, while AI systems can contribute to solving many societal challenges, certain AI systems may create risks that must be carefully addressed to avoid undesirable outcomes. To that end, it is of high significance for Vietnam to research and learn from the experiences of pioneering countries in that field of technology, such as the EU and the US, to take the right steps in developing the laws and policies AI to promote R&D and application of AI.

The EU supports a regulatory and investment-oriented approach with the twin objective of promoting the uptake of AI and of addressing the risks associated with AI technology.[3] The General Data Protection Regulation (GDPR),[4] the EU regulation on data protection and privacy, has quickly become the international standard in terms of data protection. The GDPR lays down rules relating to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system.[5] To date, nearly 120 nations have adopted the legislation

1 Revilla D. J. (2016), 'Vietnam 30 years after Doi Moi: Achievements And Challenges', *Zeitschrift für Wirtschaftsgeographie*. Retrieved from: https://www.researchgate.net/publication/309449779_Vietnam_30_years_after_Doi_Moi_Achievements_and_challenges [accessed 12 December 2022].

2 Decision No. 127/QD-TTg of the Prime Minister, dated January 26, 2021 on National Strategy for Research, Development, and Application of Artificial Intelligence Until 2030.

3 European Commission (2020), *White Paper On Artificial Intelligence - A European approach to excellence and trust*, COM (2020) 65 Final. Retrieved from: https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf [accessed 20 December 2022].

4 Regulation (EU) 2016/679 of the European Parliament and of the Council, dated 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 (General Data Protection Regulation - GDPR). Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&qid=1672629039503&from=EN [accessed 12 December 2022].

5 Art. 2.1, GDPR.

on data protection inspired with the GDPR rules.[6] In the efforts to further improve the AI ecosystem for EU, the European Commission has recently presented the AI Act,[7] which shall set out horizontal rules for the development, commodification and use of AI-driven products, services and systems within the territory of the EU.[8] The AI Act provides core artificial intelligence rules that apply to all industries.[9] It is expected that the extensive AI Act can create a new AI paradigm, which requires AI to be legally, ethically and technically robust, while respecting democratic values, human rights and the rule of law. [10]

This article aims to provide insight views of the current development of the EU law on AI system[11] and data protection by scrutinizing the interplay between both the AI Act/AI systems and the GDPR. It also analyse the implications of the EU regulations on AI and data governance in banking sector, thereby helping the legislators to assess the impacts of AI rules on data privacy and data protection in the sector that strongly supports the application of AI technology. These authors believe that these analyses can give some important suggestions for Vietnam in developing its regulations relating to the use of AI systems taking into account the liability for the use of AI systems in management and data protection.

1. GDPR'S approach toward artificial intelligence and data protection

1.1. AI system and black box

The term "Artificial intelligence" potentially covers a wide spectrum of technology but normally refers to systems that do not follow pre-programmed instructions and instead learn for themselves. This might either be using an existing data set, for example in supervised or unsupervised learning, or by prioritizing actions that lead to the best outcomes.

6 Anu B. A. (2019), *The Brussels Effect: How the European Union Rules the World*, Oxford University Press, p. 151.

7 European Council, *Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain EU Legislative Acts*, COM(2021) 206 final, dated 21 April 2021. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence. [accessed 12 December 2022].

8 *The AI Act is part of a package of digital services regulation in the EU. It forms part of a series of legislative initiatives being progressively introduced in the EU: the Digital Services Act, the Data Governance Act, the Digital Markets Act, the Data Act, the Cybersecurity Directive, the e-Privacy Regulation and the Artificial Intelligence Act.*

9 *Supra* note 7.

10 Kop M. (2021), 'EU Artificial Intelligence Act: The European Approach to AI', *Transatlantic Antitrust and IPR Development*. Retrieved from: https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai, last visited 10 December 2022 [accessed 12 December 2022].

11 *[AI system] means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*. Article 3 of the Draft AI Act. *Supra* note 7. See also OECD, 'Artificial Intelligence & Responsible Business Conduct'. Retrieved from: http://mneguidelines.oecd.org/RBC-and-artificial-intelligence.pdf [accessed 12 December 2022].

One of the implications of this behavior being learned, and not programmed, is that it may not be clear how the system reaches a decision. It operates in a "black box" is deems to be an impenetrable system.[12] Deep learning modeling is typically conducted through black box development: the algorithm takes millions of data points as inputs and correlates specific data features to produce an output. That process is largely self-directed and is generally difficult for data scientists, programmers and users to interpret.[13] When the working of software used for operation of a system cannot be assessed or reviewed easily, the errors of the system can go unnoticed, undetected until they cause serious problems with excessive damages to the involved parties. The consequences may not be limited to expenses for investigating and repairing the software errors, as sometimes the damages impossible to repair.

AI bias, for example, can be introduced to algorithms as a reflection of conscious or unconscious prejudices on the part of the developers, or they can creep in through undetected errors. In any case, the results of a biased algorithm will be skewed, potentially in a way that is offensive to people who are affected. Bias in an algorithm may come from input data when details about the dataset are unrecognized.

The AI system might work perfectly in a development environment but becomes unpredictable or unreliable in the real world. Unlike a human, the algorithm has no higher-level assessment, no god feeling. There is neither "common sense" nor "ethical" override.[14]

This creates a huge variety of legal concerns. The underlying algorithm might be making decisions that are biased or discriminatory and in breach of the broad fairness requirements of the GDPR.

1.2. Expanding the law to the "black box" of an AI machine

Given the accountability duties under the GDPR, the user of AI will not only need to ensure the machine's decision-making process is fair but also to demonstrate this is the case. This is likely to be challenging where that decision is taken in a "black box", though the use of counterfactuals and other measures may help.

Finally, there is a risk that the system will make decisions that are either discriminatory or reflect biases in the underlying dataset. This is not

12 *The black box in aviation is an extremely secure device designed to provide researchers or investigators with highly factual information about any anomalies that may have led to incidents or mishaps during a flight. The black box in AI has taken on the opposite meaning. The latest approach in Machine Learning, where there have been 'important empirical successes,' is Deep Learning, yet there are significant concerns about transparency. See* Yoshua B. (2013), 'Deep Learning of Representations: Looking Forward' in Dediu A., Martín-Vide C., Mitkov R., and Truthe B. (eds.), *Statistical Language and Speech Processing: First International Conference,* Springer, pp 1-137.

13 Wigmore I. (2019), '*Black box AI*'. Retrieved from: https://www.techtarget.com/whatis/definition/black-box-AI [accessed 12 December 2022].

14 *Ibid*.

only a potential breach of data protection law but might also raises broader ethical concerns. This reflects the common-sense expectation that important decisions, for example whether to offer someone a job or provide a mortgage, should not be entirely delegated to a machine.

Under the GDPR, this type of automated decision-making can only take place in the following situations:[15]

- *Human involvement* – If a human is involved in the decision-making process, it will not be a decision based solely on automated processing. However, that involvement would have to be meaningful and substantive. It must be more than just rubber-stamping the machine's decision.

- *Consent* – Automated decision-making is permitted where the individual has provided explicit consent. While this sounds like an attractive option, the GDPR places a very high threshold on consent and this will only be valid where the relevant decision-making process has been clearly explained and agreed to.

- *Performance of contract* – Automated decision-making is also permitted where it is necessary for the performance of a contract or to enter into a contract. An example might be carrying out credit checks on a new customer or considering whether to offer someone a job.

- *Authorized by law* – Finally, automated decision-making processing is permitted where it is authorized by law.

Even where automated decisions are permitted, AI system user must put suitable safeguards in place to protect the individual's interests. This means notifying the individual and giving them the right to a human evaluation of the decision and to contest the decision made by a machine.

1.3. Explainability

The GDPR also requires the service providers to inform individuals what information are hold about them and how it is being used. This means that if an organisation is going to use artificial intelligence to process someone's personal data, it normally needs to tell them about it.

More importantly, where automated decision-making takes place, there is a "right of explanation". Affected individuals shall be informed of the fact that automated decision-making actually happens, the significance of the automated decision-making, and how the automated decision-making operates.

The obligation boils down to providing "*meaningful information about the logic involved*".[16] This can be challenging if the algorithm is opaque. The logic used may not be easy to describe and might not even be understandable in the first place. These difficulties are recognized by regulators who do not expect

15 Art. 22 Sec. 2 and 3 GDPR.
16 Art. 15 Sec. 1 h GDPR.

organizations to provide a complex explanation of how the algorithm works, nor to disclosure of the full algorithm itself.[17]

However, AI system user should provide as full a description about the data used in the decision-making process as possible, including matters such as the main factors considered when making the decision, the source of the information and its relevance. The higher the level of autonomy of the related AI system, the more challenging it is to describe the envisaged data processing activity.

The opacity of AI systems constitutes a challenge with regards to the common view that transparency is the key prerequisite to facilitate and enable control over AI systems. It goes without saying, that the opacity issue is a roadblock to sharing information on the algorithm or the internal logic on how the data is being processed. There is a view that explainability relates to the issue on *why* a particular decision was reached, not *how*.[18] That necessitates the disclosure of different factors which have been factored in when reaching a particular output generated by the AI system. Whenever a decision is fully or partly based on the output of an AI system, one may only be able provide an explanation if the AI system is able to produce an outcome that is understandable. Some degree of explainability of an AI system is the prerequisite to provide an explanation[19] and, more importantly, to open the path to understandability.[20]

Another commentator reaches the conclusion that a right to an explanation in the GDPR seems unlikely to help find complete remedies arguing that, first, the law is restrictive on when any explanation-related right can be triggered and, second, a meaningful information about the logic of processing is unlikely to be provided by the kind of Machine Learning "explanations" computer scientists have been developing. With that, it seems preferable to utilize other tools provided by the GDPR rather than sticking to the right of explanation which leaves the individual with the burden of pursuing to pull out a meaningful version of the interior of a black box. Those

17 European Commission (2018), *Guidelines on Automated Individual Decision-Making and Profiling*, Article 29 Data Protection Working Party, WP 251 rev 01, p. 25.

18 Fink, M., & Finck, M. (2022), 'Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration', *European Law Review*, 47(3), 376-392 (384). Retrieved from: https://scholarlypublications.universiteitleiden.nl/access/item%3A3439726/view [accessed 12 December 2022].

19 *Ibid*.

20 *Information Technology Industry Council defines transparency in AI as being clear about how a specific system is built, operated and functions. One key recommendation is for policymakers to craft legislation that offers information so the public can understand the decisions a given AI system makes, while featuring the ability to review or challenge decisions*. See also Information Technology Industry, *Global Policy Principles for Enabling Transparency of AI System*. Retrieved from: https://www.itic.org/news-events/news-releases/iti-publishes-global-policy-principles-for-enabling-transparency-of-ai-systems?mkt_tok=MTM4LUVaTS0wNDIAAAGHB4gTgEuP5ZhM7enDgbcMhItHABLMvsXCz43xiTQYcruGJIK3-mL0BxQRXIfpvDMt15746vubZz4y4CGDc9nUVVQbHnF9jX4Aur4Ho7HhgLHn [accessed 12 December 2022].

alternatives boil down to focusing on the mandatory requirements for Privacy by Design and Default and data protection impact assessments as well as the opportunities for certification schemes. Those mechanisms enable individuals to be assured about algorithmic governance being in sync with the legal and regulatory requirements.[21]

Overall, different degrees of explanation detail may be necessary, depending on the individuals and the context. The appropriate approach will be the one that can clearly describe to the audience the path taken to the decision-making since the training and creation of the model.

1.4. Safeguards when using AI

None of these challenges necessarily prevents the use of artificial intelligence so long as it is used in a safe and controlled manner. Deployed properly with appropriate safeguards, artificial intelligence offers a range of potential benefits when it comes to decision-making, such as reducing the error or unconscious biases that arise in human-decision-making.

The sorts of safeguards that one might expect to see include:

- *Counterfactuals* - For example, if a loan application is rejected by an AI system it could provide the applicant not just with that rejection but also with an assessment of the minimum change needed for the application to be successful (e.g. the loan would be granted if it were for £2,000 less or the borrower's income £5,000 more). These counterfactual edge cases will provide some insight into the decision-making process.

- *Verification* – It may be possible to provide some form of human verification of the artificial intelligence's decision-making process. For simple tasks, there might be easy ways to do this. For example, a picture classification algorithm might highlight the pixels that strongly influence the classification decision.

- *Testing and output analysis* – The system should be thoroughly tested on a robust dataset. The decisions of the system should also be analyzed to ensure it is not making discriminatory or inappropriate decisions. For example, to confirm a system used to shortlist candidates for interview is not preferring female candidates over males (*or vice versa*).

- *Training* - A Machine Learning model that is deployed must be continuously trained to teach the system further correlations from incoming data, no matter how much data it is given. This means that, unless Machine Learning models continue to be trained, they cannot be expected to evolve. This is a crucial risk for the accuracy of the AI system, as its obsolescence towards reality can endanger the ability to make adjusted and fair judgements. Irrespective of the volume of (personal) data being fed in the AI system, it is

21 Lilian E. and Veale, M. (2017), 'Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For', *16 Duke Law & Technology Review* 18 (2017).

key that the underlying algorithm is constantly tested and developed.[22]

- *Ongoing sampling* – A sample of outputs from the system should be reviewed on an ongoing basis to confirm the quality of its output, particularly where the system is used in a dynamic environment.

- *Circuit breakers* – It will usually be worth adding circuit breakers to the system so that if its outputs exceed certain limits, either a warning is triggered, or the system is suspended.

Similar controls are already required under MiFID II for financial services firms carrying out algorithmic trading and high-frequency trading.

1.5. Privacy impact assessments

The interaction between artificial intelligence and the GDPR thus engages a number of relatively complex legal and technical issues that require a number of judgements.

In most cases, one will need to document this evaluation. This will be through:

- *Data protection impact assessment* – These are mandatory if the processing is "high risk" and must involve the organisation's data protection officer (if appointed). If the assessment shows that there are unmitigated high risks, the AI system user must consult its data protection regulator before rolling out that system;[23] or

- *Legitimate interest assessment* – If the processing is based on the so-called "legitimate interests test" the UK Information Commissioner[24] will expect to see that assessment documented.[25] This is a much quicker and more lightweight process than a full data protection impact assessment and can be recorded in a relatively short and informal document.

In many cases, the deployment of artificial intelligence systems will trigger the need for a full data protection impact assessment. EU guidance indicates that the use of new technology, automated decision-making and similar activities will trigger the need for a data protection impact assessment.[26] In the UK, the Information Commissioner has issued a list of activities that prima facie will require a data protection impact assessment. It specifically refers to

22 AEDP (2020), *10 misunderstandings about Machine Learning*, Joint paper from the Spanish data protection authority and the European Data Protection Supervisor. Retrieved from: https://edps.europa.eu/system/files/2022-09/22-09-20_10-misunderstandings-on-machine-learning_en.pdf [accessed 12 December 2022].

23 Art. 35 GDPR

24 UK Information Commissioner Office, Guide to the General Data Protection Regulation (GDPR), Retrieved from:

25 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/how-do-we-apply-legitimate-interests-in-practice/ [accessed 12 December 2022].

26 European Commission, *Guidelines on Data Protection Impact Assessment*, WP 248 rev 01. Retrieved from: https://ec.europa.eu/newsroom/article29/items/611236 [accessed 12 December 2022].

"Artificial intelligence, machine learning and deep learning" as a factor that may trigger the need for such an assessment.

1.6. Personal data security requirements on the use of AI

Personal data security is significant in the context of AI systems. Generally, personal data must be processed in a way that ensures appropriate levels of security against its unauthorized or unlawful processing, accidental loss, destruction or damage. However, AI may worsen the security risks and make them more difficult to control. Thus, data privacy and security requirements must be carefully regulated.

1.6.1. GDPR requirements on the use of AI

The GDPR applies to the processing of personal data in the context of an EU establishment, or when offering goods or services to, or monitoring the behavior of, individuals in the EU.[27] The GDPR applies regardless of the *means* by which personal data are processed, and therefore applies when an AI system is used to process personal data (e.g., when using an AI system to filter applications for a job vacancy).

Under the GDPR, *profiling* is the use of personal data to evaluate certain personal aspects relating to a natural person,[28] in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements.[29]

Automated decision-making is any decision made without meaningful human involvement that has legal effects on a person, or similarly significantly affects him or her. This may partially overlap with, or result from, profiling, but this is not always the case.

The GDPR imposes specific requirements on profiling and automated decision-making. The use of an AI system in relation to individuals often involves profiling, and sometimes automated decision-making. For example, when using an AI system to filter applications for a job vacancy, profiling is used to determine whether the applicant is a good fit for the vacancy. If the AI system filters out the applicants that it considers a good fit, and the other applicants are not considered for the position, this is automated decision-making towards the latter group. Their applications were removed from consideration for the position, with no meaningful human involvement.

1.6.2. Legal requirements for AI users

The GDPR imposes legal requirements on whoever uses the AI system for profiling and/or automated decision-making purposes, even if they

27 Art. 3 sec. 2 GDPR.
28 Art. 22 sec. 1 GDPR.
29 Recital 71 of the GDPR.

acquired the system from a third party. The requirements are as follow:[30]

- *Fairness*, which includes preventing individuals from being discriminated;
- *Transparency* towards individuals, including meaningful information about the logic involved in the AI system; and
- *The right to human intervention*, enabling the individual to challenge the automated decision.

1.6.3. Contractual requirements for providers

If a company acquires an AI system from a vendor, the company is often not in a position to comply with the above-mentioned requirements on its own. For example, the company may not know whether the AI is trained to prevent any discrimination or know the logic that the AI system relies on. To be able to comply with its obligations under the GDPR, a company therefore needs to rely on the vendor and will want to impose contractual obligations on the vendor to secure its cooperation and support.

2. New development of legal framework for AI under the AI Act

On 21 April 2021, the European Commission proposed the draft AI Act, which addresses the risks stemming from the various uses of AI systems. If adopted, the new regulation would be the first comprehensive regulatory scheme to focus solely on the development and use of AI. It would establish rules on the development, market placement, and use of AI systems across all sectors, all industries within the territory of the EU.[31] It would also apply to **any** providers or distributors of AI who place their services or products in the EU market.[32] Mark MacCarthy and Kenneth Propp have called the draft AI Act as "a comprehensive and thoughtful start to the legislative process in Europe [that] might prove to be the basis for trans-Atlantic cooperation."[33] It is expected that it will create a new AI paradigm, which requires AI to be legally, ethically and technically robust, while respecting democratic values, human rights and the rule of law.

2.1. What's new under the AI Act?

Compared to the GDPR, the AI Act introduces new obligations for vendors of AI systems, prohibits certain very high-risk AI systems, and introduces more specific requirements for high-risk AI systems and users thereof.

30 European Commission (2017), *Guidelines of the European Data Protection Board on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, adopted on 3 October 2017, as last Revised and Adopted on 6 February 2018. Retrieved from: https://service.betterregulation.com/document/306193 [accessed 12 December 2022].

31 *Supra* note 10.

32 Art. 2, Draft AI Act. *Supra* note 8.

33 MacCarthy M. and Propp K. (2021), Machines learn that Brussels writes the rules: The EU's new AI regulation, Brookings Institution. Retrieved from: https://www.brookings.edu/blog/techtank/2021/05/04/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/ [accessed 12 December 2022].

2.1.1. Extraterritorial scope

The AI Act defines *AI systems* as any software that, for a set of human-defined objectives, can generate outputs such as content, predictions, recommendations, or decisions influencing the environments where they interact.[34]

The AI Act applies to vendors ("providers") of AI systems as well as users of AI systems. A provider is a developer that offers an AI system on the market, whereas a user is using an AI system under its own authority. In respect of providers and users, the AI Act applies to:

- *EU and non-EU providers* that place AI systems on the EU market;
- *EU users* of AI systems; and
- Providers and users of *non-EU AI systems*, if the output of the AI system is used in the EU.

2.1.2. Prohibition on specific AI systems

Although the GDPR imposes stringent requirements on certain processing activities, it does not outright prohibit any activities. However, the AI Act explicitly prohibits a number of AI systems that are deemed too risky under any circumstances. Most of these prohibitions are limited to AI systems used by public authorities or law enforcement. The prohibited AI systems that are relevant to the private sector are those that cause physical or psychological harm to an individual by:

- Deploying subliminal techniques to distort behavior; or
- Exploiting vulnerabilities of a specific group of individuals due to their age or physical or mental disability.[35]

2.1.3. High-risk AI systems

Most of the requirements of the AI Act apply to high-risk AI systems albeit most AI systems will not be high-risk at all. The key high-risk AI systems for the private sector are AI systems used for:

- "Real-time" and "after the fact" *remote biometric identification* of individuals (e.g., facial recognition);
- *Recruitment and selection*, such as advertising job vacancies and screening or filtering applications, and evaluating candidates in the course of interviews or tests;
- *HR purposes*, such as making decisions about promotions and terminations of work-related contractual relationships, for task allocation, and for monitoring and assessing performance and behavior; and
- Evaluating the *creditworthiness* of individuals or establishing a *credit score*.

34 Art.3 AI Act.

35 See Title II Art. 5 Draft AI Act.

2.2. General requirements for high-risk AI systems

The AI Act imposes the following general requirements on high-risk AI systems:[36]

- *Establish a risk management system* and maintain it continuously throughout the lifetime of the system to identify and analyze known and foreseeable risks, estimate and evaluate such risks, and adopt suitable risk management measures;

- *Provide training, validation, and testing data*, including the relevance, representativeness, accuracy, and completeness thereof, and bias monitoring, detection, and correction, for which special categories of personal data may be used based on the substantial public interest exemption of Article 9(2)(g) GDPR;

- *Draw up technical documentation* before the AI system is placed on the market that demonstrates that the AI system complies with the AI Act;

- *Create automatic logs* to ensure a level of traceability of the system's functioning;

- *Ensure transparency* to enable the user of the system to interpret the AI system's output and use it appropriately;

- *Enable human oversight* on the AI system aimed at minimizing the risks to health, safety, or fundamental rights, by an individual who fully understands the system's capabilities and limitations and can decide not to use the system or its output in any particular situation; and

- *Ensure accuracy, robustness, and cybersecurity* to foster resilience regarding errors, faults, inconsistencies, technical faults, unauthorized use, or exploitation of vulnerabilities.

The AI Act introduces general requirements for high-risk AI systems, which are more specific than the requirements under GDPR.[37] For example, the AI Act imposes specific requirements on training, validation, and testing data in order to prevent bias and discrimination, while the GDPR merely requires that any processing of personal data is fair (including not being discriminatory). Another example is the requirement of human oversight. Although the GDPR grants individuals the right to obtain human intervention in cases of automated decision-making (as set out above), this requirement applies only to the company that makes the automated decision, and not to company that provides the related AI system.

2.2.1. Specific requirements for providers of high-risk AI systems

The GDPR applies to the processing of personal data, and not (directly) to provider of the systems that enables the processing activity. This means that companies that use third-party systems to process personal data have to

36 See Title III Chapter 2 AI Act.

37 Art. 52 AI Act: *New Transparency Obligations For Certain AI Systems*.

take on the responsibility under the GDPR since they are considered as data controllers. However, the AI Act imposes the following specific requirements on the *provider* of the high-risk AI system:[38]

- *Ensure compliance* with the above-mentioned requirements for high-risk AI systems;

- *Implement a quality management system*, including a strategy for regulatory compliance, and procedures for design, testing, validation, data management, and record keeping;

- *Address (suspected) non-conformity* by immediately taking the necessary corrective actions to (i) bring the AI system into conformity, (ii) withdraw, or (iii) recall it;

- *Notify relevant authorities about nonconformity of, or serious incidents pertaining to, the AI system* and the corrective measures taken in the countries in which the AI system has been made available;

- *Conduct conformity assessments* which can be internal or external assessments, depending on the type of high-risk AI system used;

- *Register in the AI database* before offering a high-risk AI system on the market; and

- *Conduct post-market surveillance*, by collecting and analyzing data about the performance of high-risk AI systems throughout the system's lifetime.

2.2.2. Specific requirements for users of high-risk AI systems

The AI Act imposes fewer obligations on users of high-risk AI systems than on the providers thereof, which are different from the requirements that apply under the GDPR. The AI Act requires users of high-risk AI systems to:[39]

- *Abide by the provider's instructions* on the use of the AI system, and take all technical and organizational measures indicated by the provider to address residual risks of using the high-risk AI system;

- *Ensure input data is relevant* if the user has control over such data, for example, that information on the applicant's religion is not input into an hiring AI system;

- *Monitor the operation of the system* for anomalies or irregularities;

- *Maintain log files* if the logs are under the control of the user; and

- *Notify the provider about serious incidents and malfunctioning*, and, in such a case, suspend use of the AI system.

2.3. Sanction regime under the AI Act

The AI Act provides the following penalties for non-compliance in the private sector which partly differs from the sanction regime under the GDPR:

38 See Title III Chapter 3 AI Act.
39 *Ibid*.

- Up to *EUR 30,000,000 or 6%* of the total worldwide annual turnover (whichever is higher) for non-compliance with the prohibited AI systems or the data and data governance requirements;

- Up to *EUR 20,000,000 or 4%* of the total worldwide annual turnover (whichever is higher) for non-compliance with other than the above-mentioned obligations; and

- Up to *EUR 10,000,000 or 2%* of the total worldwide annual turnover (whichever is higher) for providing incorrect, incomplete, or misleading information to competent authorities or conformity assessment entities.

2.4. What are the "war zones" between GDPR and AI systems?

2.4.1. Principle of transparency

Pursuant to Article 5 section (1)(a) GDPR personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject. Moreover, the controller shall provide the data subject with all the information laid down in Article 13 section (1) GDPR. The data subject has the right to obtain from the controller confirmation as to whether or not personal data concerning him or her is being processed (Art. 15 GDPR, right of access).

These rights of the related individual may clash with AI systems since it may remain unclear which kind of information is actually required to get the data subject to make an informed decision. The issue boils down to the question as to whether it is legally mandatory to deliver an "information overflow" vis-à-vis the individual concerned or does it suffice to convey only that kind of information he or her can easily understand.

2.4.2. Principle of purpose limitation

Article 5 section (1)(b) GDPR states that personal data shall be collected for specific, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. Any further processing shall not be permitted unless the data subject has consented to the new purpose. The reason for establishing this principle is to prevent the use of the personal data of individuals in a way or for further purposes that they might find unexpected, inappropriate, or otherwise objectionable.

The principle of purpose limitation cannot be fully aligned with AI since it is practically impossible to determine the purposes of an AI system from the outset/in the design phase. Moreover, the AI system may lay down its own purpose when processing personal data. From the data protection perspective, any data processing operation for "unknown" or hypothetical purposes which are not determined beforehand, is clearly not admissible. One possible solution is to interpret the legal term "purpose" very broadly albeit there are considerable reservations against that approach amongst data privacy pros.

2.4.3. Principle of data minimization

The principle of data minimization (Article 5 section (1)(c) GDPR) essentially means that only those data which is necessary to meet the purpose determined by the controller must be collected and processed. It partly overlaps with the principle of purpose limitation. The data minimization principle requires, in particular, to ensure that the period for which the personal information is stored is limited to a strict minimum.

In an AI environment, an enormous volume of data (not only personal data) is collected ("Big Data"). The way an AI system works, actually clashes with the data minimization principle. That principle necessitates to limit the volume of personal data being processed to a strict minimum. From the GDPR perspective, it is not a proportionate practice to increase substantially the amount of personal data in the training dataset to have only a slight improvement in the performance of the AI system. More data will not necessarily improve the performance of Machine Learning models. On the contrary, more data could bring more bias. [40]

2.4.4. Principle of data accuracy

The principle of data accuracy (Article 5 section (1)(d) GDPR) requires the data controller to keep the data accurate and up to date. The AI Act goes beyond and sets out more specific requirements for high-risk AI systems (i.e. data governance procedure) when feeding information (not only personal data) in such systems.[41] Consequently, the particular test-, training- and validation data file must be relevant, representative, complete and correct prior to making use of it.

2.5. EU Legislative Process on AI systems

According to the EU AI Act Newsletter #12 as of 2 September 2022, the rapporteurs Brando Benifei and Drago Tudorache circulated new compromise amendments.[42] On the high-risk obligations, one of the changes is that providers must immediately inform distributors and, when applicable, other actors in the value chain of any non-compliance and corrective action. When it comes to the distribution of responsibility, discussions are ongoing as to how to address the issue of allocating responsibility across the complex AI supply chain.[43]

Moreover, the center for Data Innovation argued that the EU should clarify the distinction between explainability and interpretability in the AI Act. It referred to Article 13 which states that *"High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to*

40 *Supra* note *22*.
41 Art. 10 of the AI Act.
42 Risto Uuk (2022), 'The EU AI Act Newsletter # 12', Retrieved from: https://artificialintelligenceact.substack.com/p/the-eu-ai-act-newsletter-12?utm_source=substack&utm_medium=email [accessed 12 December 2022].
43 *Ibid*.

enable users to interpret the system's output and use it appropriately". The center was challenging this, noting there are no specifics on what it means to "interpret" AI system's output nor on which technical measures a provider must take to demonstrate system compliance.[44]

The UK government published its policy paper on the future regulation of artificial intelligence (AI) on 18 July 2022 (the AI Paper), taking a less centralized approach than the EU. This follows the National AI Strategy which was published in September 2021. The AI Paper differs from the European approach by taking a pro-innovation approach.[45]

The European Commission ('the Commission') announced, on 28 September 2022, that it had issued a proposal for an AI Liability Directive ("Directive"),[46] also publishing a dedicated page with frequently asked questions (FAQs). In particular, the Commission noted that the Directive seeks to introduce, for the first time, rules specific to damages caused by AI systems, with the aim to ensure that victims of harm caused by AI technology can access reparation, in the same manner as if they were harmed under any other circumstances. To this end, the Commission continued, the proposed Directive includes two main measures:

- An alleviation of the burden of proof in relation to damages caused by AI systems, which would relieve victims from having to explain in detail how the damage was caused by a certain fault or omission; and

- The access to evidence from companies or suppliers, when dealing with high-risk AI.

Notably, the Commission clarified that both the Directive and the AI Act represent two sides of the same coin, noting that they would apply at different moments and reinforce each other. In this regard, the Commission explained that, while the AI Act aims at preventing damage, the proposed Directive lays down a safety-net for compensation in the event of damage.

Overall, the proposed Directive strives to harmonize national liability rules for AI, making it easier for victims of AI-related damage to get compensation from both providers and, under certain circumstances, users of AI systems. The new rules will ensure that victims benefit from the same standards of protection when harmed by AI products or services. The Commission´s proposal will now need to be adopted by the European Parliament and the Council.

44 *Ibid*.

45 Coleclough L. and Abercromby K. (2022), 'The UK's AI Action Plan'. Retrieved from: https://www.lexology.com/library/detail.aspx?g=2c4d1ac5-2cfb-4961-b4bd-6f0e963f94f5&utm_source=Lexology+Daily+Newsfeed&utm_medium=HTML+email+-+Body+-+General+section&utm_campaign=Lexology+subscriber+daily+feed&utm_content=Lexology+Daily+Newsfeed+2022-09-27&utm_term= [accessed 12 December 2022].

46 European Council, *Proposal for a Directive on adapting non contractual civil liability rules to artificial intelligence*. Retrieved from: https://ec.europa.eu/info/files/proposal-directive-adapting-non-contractual-civil-liability-rules-artificial-intelligence_en [accessed 12 December 2022].

3. Implication of the EU rules of the use of AI systems in the financial industry

Banks have always relied on predictions to make their decisions. Estimating the risks or rewards of making a particular loan, for example, has traditionally fallen under the purview of bankers with deep knowledge of the industry and extensive expertise. But times are changing. Today, banks realize that data science can significantly speed up these decisions with accurate and targeted predictive analytic tools. By leveraging the power of automated machine learning, banks have the potential to make data-driven decisions for products, services, and operations.

Artificial intelligence and machine learning in finance encompasses everything from chatbot assistants to fraud detection and task automation. Most banks (80%) are highly aware of the potential benefits presented by AI, according to Insider Intelligence's AI Banking report.[47] The three main channels where banks can use AI to save on costs are

- Front- and middle-office AI applications: they offer the greatest cost savings opportunity across digital banking;
- Leveraging algorithms on the front end to smooth customer identification and authentication, mimic live employees through chatbots and voice assistants, deepen customer relationships, and provide personalized insights and recommendations; and
- Leveraging AI within middle-office functions to assess risks, detect and prevent payments fraud, improve processes for anti-money laundering (AML) and perform know-your-customer (KYC) regulatory checks.

One of the most significant use cases for AI in finance is its ability to prevent fraud and cyberattacks. Consumers look for banks and other financial services that provide secure accounts, especially with online payment fraud losses expected to jump to $48 billion per year by 2023. AI has the ability to analyze and single-out irregularities in patterns that would otherwise go unnoticed by humans. Particularly, it is vital for banks offering consumer banking products to implement a proprietary algorithm to detect fraud patterns when processing credit card transactions. Details of such credit card transactions are sent to central computers which then decide whether or not the individual transaction is fraudulent.

Moreover, AI is particularly helpful in corporate finance as it can better predict and assess loan risks. For companies looking to increase their loan portfolio, AI technologies such as machine learning can help improve loan underwriting and reduce financial risk significantly.

47 Digalaki E. (2022), 'The impact of artificial intelligence in the banking sector & how AI is being used in 2022'. Retrieved from: https://www.businessinsider.com/ai-in-banking-report [accessed 12 December 2022].

While for many financial services firms, the use of AI is episodic and focused on specific use cases, an increasing number of banking leaders are taking a comprehensive approach to deploying advanced AI, and embedding it across the full lifecycle, from the front- to the back-office.

It should be noted, however, that the wide implementation of high-end technology like AI is not going to be without challenges. From the lack of credible and quality data to security issues, a number of challenges exist for banks using AI technologies:

⋆ **Data security:** One of the main challenges of AI in banking is that the amount of data collected contains sensitive information that requires additional security measures to be taken. Therefore, it is important to find the right technology partner who will provide a variety of security options to ensure your customer data is handled appropriately.

⋆ **Lack of quality data:** Banks need structured and quality data for training and validation before deploying a full-scale AI-based banking solution. Good quality data is required to ensure that the algorithm applies to real-life situations. Also, if data is not in a machine-readable format, it may lead to unexpected AI model behavior. So, banks accelerating towards the adoption of AI need to modify their data policies in order to mitigate all privacy and compliance risks.

⋆ **Lack of explainability:** AI-based systems are widely applicable in decision-making processes as they eliminate errors and save time. However, they may follow biases learned from previous cases of poor human judgement. Minor inconsistencies in AI systems do not take much time to escalate and create large-scale problems, thereby risking the bank's reputation and functioning.

To avoid catastrophic problems caused by the failure of AI system, it is important for banks to establish an appropriate level of explainability for all decisions and recommendations made by AI models. Banks need to fully understand, validate, and explain how the model makes decisions. In this context, regulatory rules governing the protection of data processed in an organization's AI system are of high significance. This would push the banks to develop appropriate AI strategy, ensuring that the system applying AI technologies operates safely and effectively in the long term. European regulatory developments in recent years on data protection under the GDPR and in particular the AI Act drafting process are moving towards this goal.

Conclusion

AI is a strategic technology that offers many benefits for society. Clearly, AI offers important efficiency and productivity gains that can strengthen the competitiveness of businesses and improve the wellbeing of peoples. Hence,

organizations using AI systems must address the risks for data subjects' privacy rights and freedoms. Data protection issues should be considered from the outset and monitored throughout the lifecycles of AI systems to ensure compliance with the fundamental human rights. For Vietnam to seize fully the opportunities that AI offers, it shall focus on developing the infrastructure and standards for data governance to ensure that the use of AI technology shall not jeopardise the fundamental rights of peoples and organisations.

The EU risk-based approach based on the pyramid of criticality with a layered enforcement mechanism of the AI Act, is an experience worth learning for Vietnam as it provides flexible and practical solutions in today`s evolving technology landscape. The AI regulations can be shaped in a way to organize and regulate AI systems by the level of risks such AI systems create. In addition, given the early stage of development of Vietnam, it is also worth to consider developing of specific regulations for particular industry on regulatory sandbox basis to ensure that industry having strong demand in the application of AI technology can deploy AI systems immediately. Regulatory sandbox regime (or experimental approach to legislation) encourages actors in a specific field to test new technologies or services in a real-life environment. Three main features characterise an experimental legal regime: temporary nature, trial-and-error approach to regulation, and collaborative involvement of stakeholders or competent authorities in the process. AI regulatory sandbox would allow private actor(s) to test their new AI systems, under the supervision of national authorities, in a flexible regulatory space. Vietnam could learn from the EU experience to develop the regulatory sandbox in data protection in certain sectors, such as banking service (since this is a sector that is well-versed new technology and able to adapt quickly to technological changes). Regulatory sandbox could be a tool to lower regulatory barriers in the sense that private entities processing personal data would not be required to comply with all applicable data protection requirements when testing the system.[48] More importantly, such development would help to mitigate the current state of lack of regulations relating to AI and data governance in Vietnam. ●

Reference

**** Books and Articles***

[1] AEDP (2020), *10 misunderstandings about Machine Learning*, Joint paper from the Spanish data protection authority and the European Data Protection Supervisor. Retrieved from: https://edps.europa.eu/system/files/2022-09/22-09-20_10-misunderstandings-on-machine-learning_en.pdf [accessed 12 December 2022]

[2] Anu B. A. (2019), *The Brussels Effect: How the European Union Rules the World*, Oxford University Press

48 Ahern D. (2021), 'Regulatory Lag, Regulatory Friction and Regulatory Transition as FinTech Disenablers: Calibrating an EU Response to the Regulatory Sandbox Phenomenon', *European Business Organization Law Review* 22(1), pp. 1-38.

[3] Ahern D. (2021), 'Regulatory Lag, Regulatory Friction and Regulatory Transition as FinTech Disenablers: Calibrating an EU Response to the Regulatory Sandbox Phenomenon', *European Business Organization Law Review* 22(1)

[4] Coleclough L. and Abercromby K. (2022), 'The UK's AI Action Plan'. Retrieved from: https://www.lexology.com/library/detail.aspx?g=2c4d1ac5-2cfb-4961-b4bd-6f0e963f94f5&utm_source=Lexology+Daily+Newsfeed&utm_medium=HTML+email+-+Body+-+General+section&utm_campaign=Lexology+subscriber+daily+feed&utm_content=Lexology+Daily+Newsfeed+2022-09-27&utm_term= [accessed 12 December 2022]

[5] Digalaki E. (2022), 'The impact of artificial intelligence in the banking sector & how AI is being used in 2022'. Retrieved from: https://www.businessinsider.com/ai-in-banking-report [accessed 12 December 2022]

[6] Fink, M., & Finck, M. (2022), 'Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration', *European Law Review*, 47(3), 376-392 (384). Retrieved from: https://scholarlypublications.universiteitleiden.nl/access/item%3A3439726/view [accessed 12 December 2022]

[7] Information Technology Industry*, Global Policy Principles for Enabling Transparency of AI System*. Retrieved from: https://www.itic.org/news-events/news-releases/iti-publishes-global-policy-principles-for-enabling-transparency-of-ai-systems?mkt_tok=MTM4LUVaTS0wNDIAAAGHB4gTgEuP5ZhM7enDgbcMhItHABLMvsXCz43xiTQYcruGJIK3-mL0BxQRXIfpvDMt15746vubZz4y4CGDc9nUVVQbHnF9jX4Aur4Ho7HhgLHn [accessed 12 December 2022]

[8] Kop M. (2021), 'EU Artificial Intelligence Act: The European Approach to AI'*, Transatlantic Antitrust and IPR Development*. Retrieved from: https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai, last visited 10 December 2022 [accessed 12 December 2022]

[9] Lilian E. and Veale, M. (2017), 'Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For', *16 Duke Law & Technology Review* 18 (2017)

[10] MacCarthy M. and Propp K. (2021), Machines learn that Brussels writes the rules: The EU's new AI regulation, Brookings Institution. Retrieved from: https://www.brookings.edu/blog/techtank/2021/05/04/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/ [accessed 12 December 2022]

[11] OECD, 'Artificial Intelligence & Responsible Business Conduct'. Retrieved from: http://mneguidelines.oecd.org/RBC-and-artificial-intelligence.pdf [accessed 12 December 2022]

[12] Revilla D. J. (2016), 'Vietnam 30 years after Doi Moi: Achievements And Challenges', *Zeitschrift für Wirtschaftsgeographie*. Retrieved from: https://www.researchgate.net/publication/309449779_Vietnam_30_years_after_Doi_Moi_Achievements_and_challenges [accessed 12 December 2022]

[13] Yoshua B. (2013), 'Deep Learning of Representations: Looking Forward' in Dediu A., Martín-Vide C., Mitkov R., and Truthe B. (eds.), *Statistical Language and Speech Processing: First International Conference,* Springer

[14] Wigmore I. (2019), '*Black box AI*'. Retrieved from: https://www.techtarget.com/whatis/definition/black-box-AI [accessed 12 December 2022]

**** Laws and Regulations***

[15] Decision No. 127/QD-TTg of the Prime Minister, dated January 26, 2021 on National Strategy for Research, Development, and Application of Artificial Intelligence Until 2030

[16] Regulation (EU) 2016/679 of the European Parliament and of the Council, dated 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 (General Data Protection Regulation - GDPR). Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&qid=1672629039503&from=EN [accessed 12 December 2022]