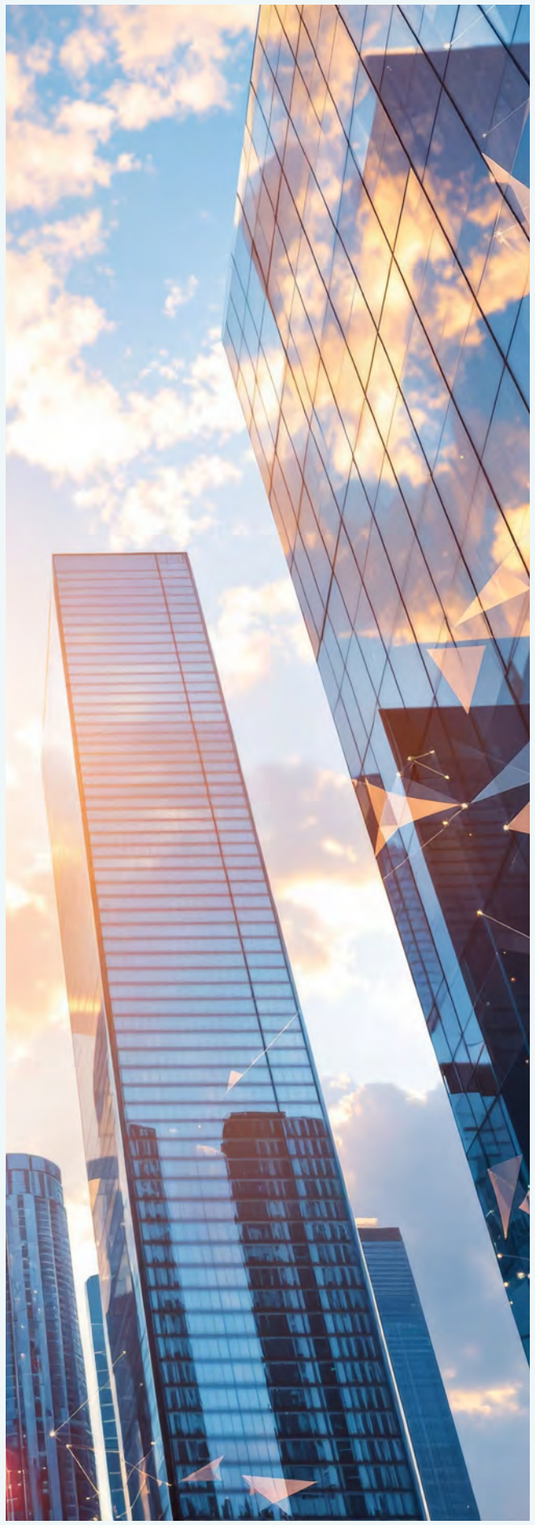
WilmerHale's Guide to AI and GDPR

A Road Map to Compliance by Design

September 2025

WILMERHALE



The rise of AI and its widespread availability offers significant growth opportunities for businesses. However, it necessitates a robust governance framework to ensure compliance with regulatory requirements, especially under the EU AI Act (see our Guide to the AI Act) and the EU GDPR.

The reason GDPR compliance is so important is that (personal) data is a key pillar of AI. For AI to function effectively, it requires good-quality and abundant data so that it can be trained to identify patterns and relationships. Additional personal data is often gathered during deployment and incorporated into AI to assist with individual decision-making.

This guide discusses GDPR compliance throughout the AI development lifecycle and when using AI.

Data Protection by Design

GDPR compliance plays a key role throughout the AI development lifecycle, starting from the very first stages. This reflects one of the key requirements and guiding principles of the GDPR, called "data protection by design" (Article 25 GDPR). Businesses are required to implement appropriate technical and organizational measures, such as pseudonymisation, both at the determination stage of processing methods and during the processing itself. These measures should aim to implement data protection principles, such as data minimisation, and integrate necessary safeguards into the processing to ensure GDPR compliance and protect individuals' data protection rights.

AI Development Lifecycle

The AI development lifecycle encompasses four distinct phases: planning, design, development and deployment. In this context, in accordance with the terminology of the EU AI Act, we will refer to both AI models and AI systems:

- **AI models** are a component of an AI system and are the engines that drive the functionality of AI systems. AI models require the addition of further components, such as a user interface, to become AI systems.
- **AI systems** present two characteristics: (1) they operate with varying levels of autonomy and (2) they infer from the input they receive how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.

Table of Contents

List of Abbreviations

AI	*Artificial Intelligence*
AI Act	*European Union's Artificial Intelligence Act*
CJEU	*Court of Justice of the European Union*
DPIA	*Data Protection Impact Assessment*
EDPB	*European Data Protection Board*
EDPB Opinion on AI Models	*Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models (adopted on 17 December 2024)*
GDPR	*European Union General Data Protection Regulation*
SMPC	*Secure Multiparty Computation*

How We Can Help

WilmerHale has a leading practice in EU law and regulation, advising clients on high-profile matters in both established and emerging market sectors across a wide variety of industries. With around 1,100 lawyers located throughout 12 offices in the United States, Europe and the United Kingdom, we offer a global perspective on EU law issues and offer single-team transatlantic and Europe-wide services. We practice at the very top of the legal profession and offer a cutting-edge blend of capabilities that enables us to handle cases of any size and complexity.

Our European offices in Brussels, Frankfurt, Berlin and London are best known for high-quality regulatory work before authorities and appellate work before EU courts. Clients entrust us with complex cases because of our expertise, reliability, responsiveness, precision and reputation with authorities. Our European team is involved in a huge number of cases in various areas of EU law, including several major data protection law cases setting breakthrough principles. In addition, many of our lawyers are qualified in several jurisdictions across the European Union, its neighbouring countries, and the United States and can handle the most complex cases requiring native-speaker proficiency in multiple languages.

Our European team works seamlessly with our US AI and Cybersecurity and Privacy teams, leveraging our combined legal expertise to provide comprehensive, cross-border support on data protection and AI-related matters. This close collaboration ensures that our clients benefit from globally informed legal strategies.

For more information on this guide or other AI or data-related matters, please contact one of the authors.

Dr. Martin Braun
Partner
Frankfurt/Brussels

Anne Vallery
Partner-in-Charge
Brussels

Itsiq Benizri
Counsel
Brussels

1.

First Phase of the AI Development Lifecycle: Planning

The first phase of the AI development lifecycle involves understanding the business problem, defining objectives and requirements, and developing a solid AI governance structure to ensure regulatory compliance. During this phase, it is essential to determine the scope of (personal) data needed and identify any constraints related to such data, with a focus on the availability of the relevant datasets.

In this context, key GDPR compliance considerations involve evaluating whether the data is personal data, ensuring the processing of the data has a valid legal basis, and verifying that the processing respects the principle of purpose limitation, including with regard to other key principles under the GDPR.

Personal Data

The GDPR only applies to personal data, i.e., any information relating to a natural person that is or can be identified, directly or indirectly. A key question, therefore, is whether AI input or output data constitutes personal data.

- **Input data** is information provided to or directly obtained by an AI system, based on which the system generates an output.
- **Output data** varies depending on the type of AI model and its intended usage. There are three major sorts of outputs: prediction, recommendation, and classification.

The European Data Protection Board (EDPB), the umbrella group of the EU's data protection authorities, issued a nonbinding Opinion on AI Models in December 2024. In the opinion, the EDPB considered whether and how AI models trained with personal data can be deemed anonymous. The EDPB identified two scenarios:

- **The AI model is designed to provide personal data.** When an AI model is specifically designed to provide personal data regarding individuals whose personal data was used to train the model or in some way to make such data available, it cannot be regarded as anonymous and the GDPR necessarily applies. According to the EDPB,

examples of such AI models include a generative model fine-tuned on the voice recordings of an individual to mimic their voice or a model designed to reply with personal data from the training when prompted for information regarding a specific person.

— **The AI model is *not* designed to provide personal data.** The EDPB considers that, even when an AI model has not been designed to produce personal data from the training data, it is still possible that personal data from the training dataset remains absorbed in the parameters of the model and can be extracted from that model. Whether the outputs of such AI models can be considered anonymous should be determined on a case-by-case basis. The EDPB appears to agree that an AI model may be anonymous, although it considers such a scenario highly unlikely. According to the EDPB, an AI model can only be anonymous provided it meets the following conditions:

- The likelihood that individuals whose data was used to build the model may be identified (directly or indirectly) is insignificant; and
- The likelihood of obtaining, intentionally or not, such personal data from queries is insignificant too.

The EDPB considers that examining whether these conditions are met must take into account the Article 29 Working Party's Guidance on Anonymisation. This guidance treats pseudonymisation merely as a security measure. However, in ***SRB*** v ***EDPS***, the Court of Justice of the European Union (CJEU) held that pseudonymised data should not be regarded as personal data in *all* cases and for *every* person (see Chapter 2).

More fundamentally, the EDPB considers that determining whether the above conditions are met must take into account whether the risk of identification has been assessed, considering all the means reasonably likely to be used to identify individuals (Recital 26 GDPR). According to the EDPB, the determination of those means should be based on objective factors, such as:

- The characteristics of the training data (e.g., the uniqueness of the records in the training data, precision of the information, aggregation, and randomization, and how these affect the vulnerability to identification), the AI model, and the training procedure.
- The context in which the AI model is released and/or processed, with contextual elements including measures such as legal safeguards and limiting access only to some persons.
- The additional information that would allow the identification and may be available to the given person.
- The costs and amount of time that the person would need to expend to obtain such additional information.
- The technology available at the time of the processing, and technological developments.

The EDPB Opinion on AI Models provides a non-exhaustive and non-prescriptive list of possible elements that may be considered when assessing AI's anonymity. These include the steps controllers take in the design stage to minimise or stop the gathering of training-related personal data and make it less identifiable, AI model testing and resistance to attacks, and documentation regarding processing operations, including anonymisation (see Chapter 2).

Legal Basis

Under the GDPR, the processing of personal data is only lawful if the controller can demonstrate a valid legal basis. The most relevant legal bases for AI under the GDPR are consent and legitimate interests. According to the EDPB, the development and deployment phases entail different processing activities that call for different legal bases and should be evaluated individually.

— **Consent.** Valid consent is often difficult to obtain because it must be individual, specific, informed, unambiguous and provided by a clear affirmative action. These conditions are generally interpreted restrictively. In addition, consent can be withdrawn at any time, and it should be as easy to withdraw consent as it is to give it.

— **Legitimate interests.** Personal data may be processed if the processing is necessary to pursue a legitimate interest and such interest is not overridden by the interests or fundamental rights and freedoms of the individuals concerned. Legitimate interests may only be relied on provided the following three-step test is satisfied, and this test must be assessed on a case-by-case basis:

- **Legitimate interest.** The processing must pursue a legitimate interest. An interest is considered legitimate if it is lawful, clearly and precisely articulated, and real and present (i.e., not hypothetical). For example, the EDPB considers that the use of a chatbot to assist users and the use of AI to improve cyber threat detection may be legitimate interests.
- **Necessity.** The processing must be necessary to pursue the legitimate interest in question. The EDPB sets a very high bar for necessity, as it considers that the assessment must evaluate the appropriate volume of personal data involved to determine whether the processing is proportionate to pursue the legitimate interest, but also whether there are less intrusive alternatives to achieve it in accordance with the data minimisation principle. In other words, the processing of personal data is not necessary if the legitimate interest can be pursued through an AI model that does not entail such processing. This is obviously a very restrictive approach.
- **Balancing test.** The legitimate interest must not be overridden by the interests or fundamental rights and freedoms of the individuals concerned. This step consists of identifying and describing the different opposing rights and interests at stake. The interests of the individuals concerned may include, for example, their interest in retaining control over their personal data, financial interests (e.g., where an AI model is used by an individual to generate revenues), personal benefits (e.g., where the individual is using AI to improve accessibility to services), or socioeconomic interests (e.g., AI that improves access to healthcare or education). Opposing interests would typically include the AI developer's fundamental right to conduct business.

The impact of the processing on individuals may be influenced by the nature of the data processed by the models (e.g., financial or location data may be particularly sensitive), the context of the processing (e.g., whether personal data is combined with other datasets, what is the overall volume of data and number of individuals affected, and whether they are vulnerable), and its consequences (e.g., violation of fundamental rights, damage, or discrimination). Importantly, the analysis of such possible consequences

must take into account the likelihood of these consequences materializing, especially considering the measures in place and the circumstances of the case.

Individuals' reasonable expectations also play a key role in the balancing test. The assessment of such expectations must take into account various criteria, such as the information provided to the individuals concerned and the wider context of the processing, including whether or not the personal information was accessible to the public, the type of relationship with the company processing personal data, the type of service, the context and source of the data collection, the possible future applications of the model, and whether people are genuinely aware that their personal data is online.

If the balancing exercise indicates that there are negative impacts from the processing on individuals, mitigation measures may tip the balance in favour of the AI developer. These steps may be technical in character (e.g., data minimisation, pseudonymisation, or the use of synthetic data—see Chapter 2), facilitate the exercise of human rights (e.g., offer an unconditional opt-out or a right to erasure that is more generous than the one enshrined in the GDPR), or improve transparency (provide extensive information to individuals, including through email campaigns or by using FAQs, graphic visualizations, and transparency labels).

Purpose Limitation

As discussed above, the planning phase involves understanding the business problem and defining objectives of the AI model or system to be developed. This is key for GDPR compliance because the GDPR requires that personal data only be collected for specified, explicit, and legitimate purposes, and that it not be further processed in a manner incompatible with those purposes. This is also important because compliance with other core GDPR principles requires a solid understanding of the purpose of AI development.

- **Transparency.** The purpose of the processing must be communicated to the individuals concerned.
- **Data minimisation.** The processing must be limited to what is necessary in relation to the purpose of the processing.
- **Accuracy.** Every reasonable step must be taken to ensure that personal data that is inaccurate with regard to the purpose for which it is processed, is erased or rectified without delay.
- **Storage limitation.** Personal data must be kept for no longer than is necessary for the purpose for which it is processed. This entails laying down protocols for the safe disposal of data, setting precise retention periods (carefully determined based on the specific needs of the AI model), and stating the need for data retention.

Data Protection Impact Assessment (DPIA)

The GDPR requires a DPIA prior to the processing when the processing is likely to result in a high risk to the rights and freedoms of individuals. In this context, the nature, scope, context, and purposes of the processing must be taken into account.

According to a recent report commissioned by the EDPB on large language models, examples of common scenarios that may require a DPIA include:

- The use of new technologies that could introduce privacy risks.
- Large-scale monitoring of publicly accessible spaces (e.g., video surveillance).
- Processing sensitive data categories such as racial or ethnic origin, political opinions, religious beliefs, genetic data, biometric data or health information.
- Automated decision-making that has legal or similarly significant effects on individuals.
- Processing children's data or any data where a breach could lead to physical harm.

Even when a DPIA is not legally required, conducting one can be prudent for best practices in AI projects. It allows organizations to preemptively address potential data protection risks, assess the impact of their solutions, and demonstrate accountability.

2.

Second Phase of the AI Development Lifecycle: Design

The second phase of the AI development lifecycle involves implementing a data strategy, focusing on data gathering and addressing potential data quality issues. It also includes converting raw data into valuable information, anonymising and minimising personal data, and implementing privacy-enhancing technologies. In this phase, key issues for GDPR compliance include data collection; data preparation (including training methodology); measures regarding outputs of the AI model; and the model's or system's architecture.

Data Collection

For AI development, (personal) data can be collected either from first-party or third-party sources.

- **First-party data** refers to personal data directly collected from the individuals concerned.
- **Third-party data** refers to personal data collected from a third party, for example, from a data broker or collected via web scraping, a commonly used technique for collecting information from publicly available online sources.

GDPR compliance requires a careful assessment of the selection of sources used to train the AI model. According to the EDPB Opinion on AI Models, this includes an evaluation of "*any steps taken to avoid or limit the collection of personal data, including, among other things, (i) the appropriateness of the selection criteria; (ii) the relevance and adequacy of the chosen sources considering the intended purpose(s); and (iii) whether inappropriate sources have been excluded.*" Typically, web scraping can be configured to ensure that specific data categories are not collected or that certain sources, such as public social media profiles, are excluded from data collection.

Data Preparation

The preparation of data for the training phase is key to GDPR compliance. This requires, according to the EDPB, careful assessment of anonymisation and pseudonymisation techniques, with consideration for minimisation and accuracy principles. These aspects are also important when choosing an AI training methodology.

- **Anonymisation.** Anonymous data is not subject to the GDPR, so anonymising personal data for AI training purposes is a good way to limit the scope of application of the GDPR (see Chapter 1). The standard for anonymising personal data is fairly high and is the subject of complex case law, especially in *Breyer* and *SRB v EDPS*. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used to identify an individual. This requires taking into account all objective factors, such as the costs of and the amount of time required for identification; the available technology at the time of the processing and technological developments (Recital 26 GDPR). The EDPB considers that AI models may be anonymous, although that is highly unlikely in its opinion (see Chapter 1).

- **Synthetic data.** An alternative to collecting and anonymising personal data can be the use of synthetic data, which avoids the complexities associated with meeting the legal standard for anonymisation. Synthetic data is based on artificial data points engineered to serve as direct substitutes for real personal data in various downstream applications. AI models learn the patterns and statistical attributes of the original data and can then be used to re-create new, entirely made-up datasets. These synthetic datasets "look and feel" like the original data and contain all the statistical information but none of the personally identifiable information.

- **Pseudonymisation.** Pseudonymisation is also a good way to mitigate GDPR compliance risks. It is one of the measures identified in Article 25 GDPR under the data protection by design approach. Pseudonymisation should be implemented taking into account the current technology, the implementation cost, and the nature, scope, context, and purposes of processing. The risks to the rights and freedoms of individuals, with varying likelihood and severity, must also be considered. In *SRB/EDPS*, the CJEU held that pseudonymised data does not constitute personal data if the pseudonymisation effectively prevents anyone other than the controller who performed the pseudonymisation from identifying the data subject, so that, *for those others*, the data subject is not identifiable. However, pseudonymous data is likely to remain personal data from the perspective of that controller. In any event, pseudonymising data helps mitigate risks, such as unauthorised access to the personal data in question. Pseudonymisation may also be a mitigating measure that may tip the balance in favour of the AI developer when relying on legitimate interests as a legal basis for the processing of personal data (see Chapter 1).

- **Minimisation.** Personal data must be adequate, relevant, and limited to what is necessary in relation to the purposes for which it is processed. This therefore requires a careful assessment of the personal data processed to determine whether it is necessary for AI development. AI models must be tested to prevent unintentional data memorisation and reduce the risk of accidentally disclosing personal data.

- **Accuracy.** Personal data must be accurate and, where necessary, kept up to date. Every reasonable step must be taken to ensure that personal data that is inaccurate, having regard to the purposes for which it is processed, is erased or rectified without delay. Data accuracy is key both for input and output data. Inaccurate personal input data is not compliant with the GDPR and will lead to inaccurate output data. The GDPR transparency principle requires informing individuals about the accuracy limits of personal data generated by AI. The AI Act requires that high-risk AI systems be designed in such a way that they achieve an appropriate level of accuracy, which must be declared in the instructions for use of the AI system in question.

Measures Regarding Outputs

Generative AI trained on personal data might unintentionally reveal some of such data when prompted. If the AI model lacks safeguards such as response filtering or differential privacy, a user could extract personal information by crafting specific queries. It is therefore critical to adopt measures to lower the likelihood of obtaining personal data related to training data from queries.

Architecture Design

In the design phase, AI engineers select the prepared data and the most suitable algorithms and techniques for the problem they are trying to solve. The architecture design should also include mechanisms for human oversight and intervention under the GDPR and the AI Act. This is quite challenging given that black-box AI models currently make up a substantial portion of the most sophisticated machine learning models on the market. These AI models are built to analyse data autonomously and in a manner that is frequently challenging to decipher from the outside. Although users can view the inputs and outputs of the system, they are unable to observe the internal workings of the AI tool that generates those outputs. Naturally, this makes it more challenging to transparently convey the intricacy of the analytical procedures used to the affected individuals.

- **GDPR and automated individual decision-making.** Save in limited exceptions, the GDPR gives data subjects the right not to be subject to decisions based solely on automated processing which produce legal effects on them or similarly significantly affect them. This right includes the right for the individuals concerned to obtain human intervention and express their point of view to contest these decisions. Thus, when designing AI, it is important to foresee the possibility of human intervention to comply with this provision. In addition, individuals must be provided with meaningful information about the logic involved in the automated individual decision-making.

  In *Dun & Bradstreet*, the CJEU clarified that this entails an obligation to explain by means of relevant information and in a concise, transparent, intelligible, and easily accessible form, the procedure and principles applied to use personal data to obtain a specific result. The mere communication of a complex mathematical formula or algorithm is not sufficient. The explanation offered must help the data subject understand and challenge the automated decision. If disclosing such information may entail the disclosure of trade secrets, the company in question must provide the relevant information to the court or supervisory authority, which will determine on a case-by-case basis whether and what information should be supplied to the data subject.

- **AI Act and Human Oversight for High-risk AI.** Under the AI Act, high-risk AI systems must be designed and developed in such a way that they can be effectively overseen by humans (see here). Human oversight must aim to prevent or minimise the risks to health, safety or fundamental rights—including the right to the protection of personal data—that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse. The oversight measures must be commensurate with the risks, level of autonomy and context of use.

3.

Third Phase of the AI Development Lifecycle: Development

The third phase of the AI development lifecycle involves building the AI model, defining its features, and transforming data into a useful representation to improve the model's performance and boost its explainability. AI training then enables algorithms to learn from the prepared dataset. This is when the model develops and enhances its capacity to make predictions by learning patterns from the data. Validation and testing further ensure model performance and generalization.

In this context, particular attention must be paid to individuals' GDPR rights and data security, which are key aspects of GDPR compliance and highly relevant for the AI development process.

Right to Information

The GDPR requires that individuals be given specific information so that they can exercise their GDPR rights.

The information to be shared varies depending on whether the personal data was collected directly from the individual concerned (Article 13 GDPR) or from another source (Article 14 GDPR).

- **If the personal data was collected directly from the individual concerned,** the information must be provided when it is obtained.
- **If the personal data was collected from another source,** the information must be shared within a reasonable period after the data is obtained, but no later than one month, taking into account the specific circumstances of the processing. More specific rules apply to intended disclosure to another recipient or communication with the individual concerned.

Any information shared must always be provided in a concise, transparent, intelligible, and easily accessible form, using clear and plain language.

Generally, individuals must know what personal data is processed, why, for how long, with whom it is shared, and their GDPR rights, including but not limited to the rights of access, rectification and erasure, and restriction, and the right to object. AI models (and systems) must be built in such a way that they can adapt if individuals exercise these rights.

Right of Access

The GDPR gives individuals the right to obtain confirmation as to whether their personal data has been processed and, when this is the case, access to such data and the information listed in Article 15 GDPR, including, but not limited to, the purposes of the processing, the categories of personal data concerned, the data recipients, and data transfers.

The right to obtain a copy of personal data can be very challenging as it involves ensuring that the rights and freedoms of others are not affected. Granting access therefore requires appropriate safeguards, including a fair amount of anonymisation.

Rights to Rectification and Erasure

Individuals have the right to obtain the rectification of inaccurate personal data concerning them. If AI includes inaccurate personal data, the individual concerned may ask that it be rectified or completed.

Accuracy has a different meaning in the context of AI development than it does in the GDPR. Although it serves different purposes in these contexts, accuracy in the GDPR and AI development impact each other.

- **Accuracy in AI.** In AI development, accuracy refers to the performance of an AI model in correctly predicting or classifying data. It is a measure of how well the model's outputs match the true values or labels in the dataset. High accuracy indicates that the AI model is reliable and effective in its tasks, such as image recognition, natural language processing, or predictive analytics.
- **Accuracy in the GDPR.** The GDPR requires that personal data is accurate and kept up to date, and that every reasonable step is taken to ensure that inaccurate personal data is erased or rectified without delay. The GDPR's focus on accuracy is aimed at protecting individuals' rights.
- **Relationship.** All personal data, whether it is an output of an AI system or information about an individual as an input, is subject to the accuracy principle. The accuracy of the output depends on the accuracy of the input. Therefore, when it comes to personal data, the model's or system's performance is inherently linked to the GDPR accuracy principle.
- **Fairness.** A separate but equally important issue is whether AI generates harmful content due to the information it ingests. For example, AI trained on data reflecting gender inequalities can generate results that discriminate against individuals based on their gender. According to the EDPB, the GDPR fairness principle requires that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected, or misleading to the individual concerned. To address the risks of bias and discrimination, it is possible to alter the learning procedure, alter the data by adding or removing data concerning underrepresented or overrepresented demographic groupings to balance the training data, or alter the model after it has been trained.

In limited circumstances, individuals have the right to obtain the erasure of their personal data. This typically applies when the data in question has been processed unlawfully or is no longer necessary for the purposes of the processing. Data erasure can disrupt the training and performance of AI models that rely on large datasets. Removing data can lead to gaps, reduce the model's accuracy, and necessitate retraining with updated datasets.

Rights of Restriction and to Object

Individuals have the right to obtain the restriction of processing while they review the accuracy of their personal data, or if the processing is unlawful or no longer necessary, or the individuals concerned have objected to the processing of their personal data based on legitimate interests. The right of restriction and the right to object may have an enormous impact on the processing of personal data for building AI models.

- **If the processing has been restricted**, except for storage purposes, the data in question may only be processed with the individual's consent or for limited purposes, such as the establishment, exercise, or defence of legal claims, and the protection of other persons' rights.
- **If the processing has been objected to,** the personal data can no longer be processed unless compelling legitimate grounds for the processing override the interests, rights, and freedoms of the individual concerned, or for the establishment, exercise, or defence of legal claims.

Security

The GDPR requires implementing appropriate technical and organizational measures to ensure a level of security appropriate to the risk, especially regarding unlawful destruction, loss, alteration, and unauthorised disclosure of or access to personal data.

It is essential to address the risks posed by potential threats that could result in the exposure of personal data processed during the AI training phase. Typical risks include model inversion, membership inference, and attribute inference.

- **Model inversion attacks** involve using the output of an AI model to infer the input.
- **Membership inference attacks** consist of determining whether a specific data point (also called a target sample) was part of the training dataset.
- **Attribute inference attacks** involve attempting to extract information about the sample from the target model. This assumes that the attacker has partial knowledge of a sample in the training set.

These threats can be mitigated using privacy-enhancing technologies, such as the following:

- **Differential privacy** works by adding random noise to the data, preventing attackers from identifying individuals while allowing useful insight to be drawn from the dataset.
- **Federated learning** allows different parties to train AI models on their own information. They then combine identified patterns into a global model without having to share any training information with each other. This helps minimise the risk arising from data breaches, as no personal data is held together in a central location.
- **Synthetic data** is artificial data generated by data synthesis algorithms to reduce the amount of personal data processed (see Chapter 2).
- **Homomorphic encryption** allows computations to be performed on encrypted information without first decrypting it. This helps minimise the risk from data breaches because personal data remains encrypted at rest, in transit and during computation.
- **Secure multiparty computation** (SMPC) allows different parties to jointly process their combined information without any party needing to share all of its information. SMPC helps minimise the risk from personal data breaches since the shared information is not stored together.

4.

Fourth Phase of the AI Development Lifecycle: Deployment

The fourth phase of the AI development lifecycle involves making AI accessible for real-world use, tracking performance, addressing drifts, and adjusting through monitoring and maintenance.

Implementing data protection by design establishes a strong foundation for GDPR compliance, but it is not sufficient on its own. GDPR compliance is an ongoing process that necessitates continuous monitoring and appropriate processes throughout the lifespan of the AI model or system to ensure that all the issues discussed in the previous chapters remain properly addressed at all times.

Monitoring and Processes

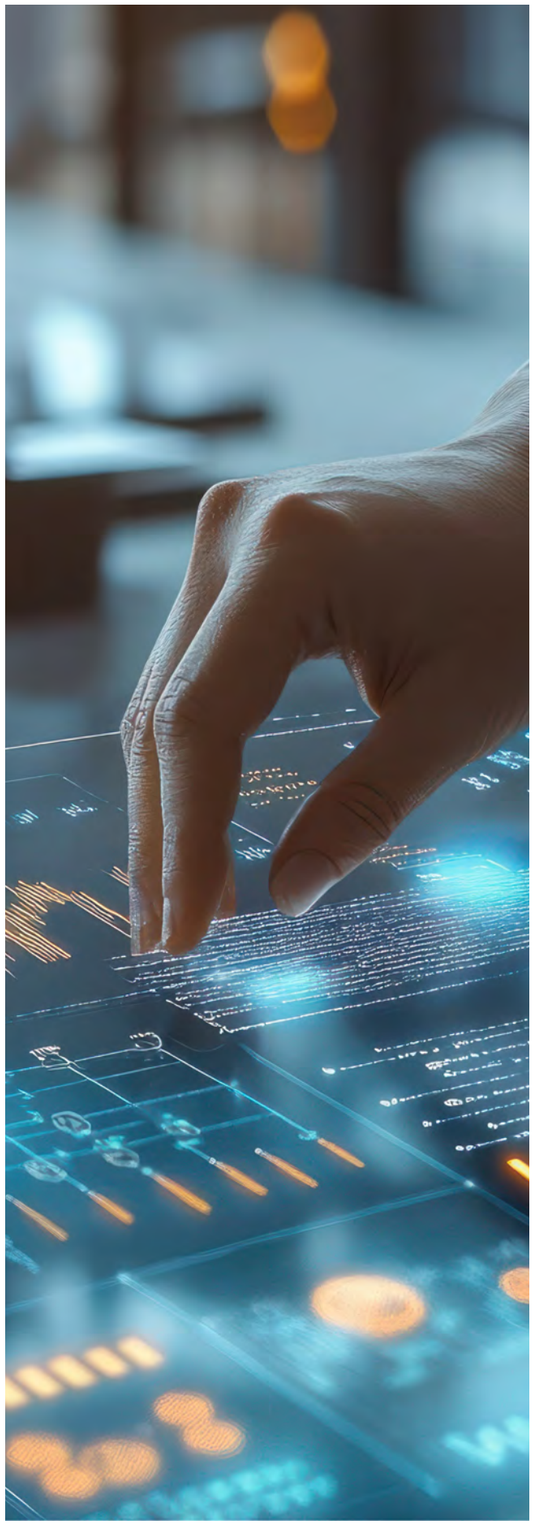
- **Monitoring.** Once an AI model or system is deployed, continuous monitoring is crucial to ensure it maintains strong performance over time, as well as GDPR compliance. By analysing key metrics and incorporating user feedback, the model's predictions should be regularly evaluated. A drop in accuracy or performance indicates that updates or retraining may be necessary, effectively closing the loop in the AI lifecycle. This continuous evaluation is vital for the model to remain adaptable and accurate in its real-time application.
- **Processes.** Ensuring appropriate processes is particularly important to comply with GDPR requirements concerning individuals' rights and notification of security breaches.

Individuals' Rights

It is essential to establish processes that address individuals' requests for information, access to their personal data, portability, rectification, erasure, restriction, and the right to object (see Chapter 3). These rights are applicable throughout the entire lifecycle of an AI system, encompassing both the personal data used in training datasets and the data processed during the system's operational phase.

Security

- **Ensuring continued security.** Another key measure that can be used to ensure ongoing GDPR compliance is shoring up the continued security of AI. Malicious actors may attempt prompt injection attacks, using harmful prompts to bypass safeguards, gain unauthorised access, extract personal data, or manipulate outputs. Additionally, attackers might design inputs specifically to capture the model's responses, gradually building a dataset of input-output pairs to train a replica model, essentially copying the original's functionality (see Chapter 3). To prevent such threats, it is important to understand how the model is being used or misused—even if that is different from what was originally expected—and align that usage with established assessment frameworks to ensure safe and responsible operation (see Chapter 3).
- **Notifying authorities of security breaches.** It is imperative to have processes in place to notify relevant authorities of security breaches without undue delay and, where feasible, no later than 72 hours after having become aware of the breach. No notification is required where the personal data breach is unlikely to result in a risk to the rights and freedoms of the individuals concerned. If that is not the case, the breach must also be communicated to the individuals concerned without undue delay.



5.

Using AI

The GDPR is applicable not only to companies that are developing AI but also to those using it. This chapter focuses on organizations acting as controllers under the GDPR, meaning entities that determine the purposes and means of processing personal data. For instance, companies are considered controllers when they employ an AI large language model to analyse employee records or generate work products that include personal data.

Joint Controllership or Controller-Processor Relationships

If the AI developer and the company using its AI solution collaboratively determine the purposes and methods of processing personal data in connection with such a solution, they are considered joint controllers. Consequently, they must enter into an arrangement to establish their respective obligations under the GDPR.

If, however, the AI developer acts as a processor, meaning it processes personal data on behalf of the company using AI (the controller), they are required to enter into a controller-processor agreement. This agreement must outline the subject matter and duration of the processing, the nature and purpose of the processing, the types of personal data involved, the categories of individuals concerned, and the obligations and rights of the parties.

- **Joint Controllership Arrangement.** The AI developer will rarely be a joint controller because it will rarely *jointly* determine the purposes of processing operations with companies using AI.
- **Controller-Processor Agreement.** AI developers typically handle personal data for companies utilizing their AI solutions. This is generally applicable to all software-as-a-service offerings. Companies using such AI solutions must have a controller-processor agreement in place and verify that the processor guarantees the implementation of appropriate technical and organizational measures to ensure GDPR compliance. This especially includes measures to ensure an appropriate level of security and compliance with the GDPR requirements for transfers of personal data outside the EU.

Data Input

Companies using AI must ensure that they comply with the GDPR when inputting personal data into AI systems. Key points for consideration are as follows:

- **Awareness.** Companies using AI must prioritise internal awareness and provide comprehensive training to all relevant staff. Compliance with GDPR cannot be achieved without adequately trained employees.
- **Purpose limitation.** Companies using AI should clearly define the purposes for which personal data is processed. AI should not be used for purposes not permitted by the company's policy.
- **Data minimisation.** Companies should limit the amount of personal data included in AI systems. It is best to provide anonymous data as input unless personal data is necessary, in which case it should be limited to what is indeed necessary. Companies need to exercise caution with freely accessible large language models, as any information provided to such systems will generally be shared with the developer of that system. Companies may therefore consider prohibiting the use of these tools or ensuring that no personal data is input into them.

Data Output

Companies using AI also need to ensure that the system's output complies with the GDPR.

- **Accuracy.** It is essential for companies using AI to verify that any personal data generated as output, or any personal data provided by the company based on such output, is accurate. As AI-generated data outputs might not always be accurate, reviewing and addressing potential inaccuracies is crucial. This process also helps mitigate any possible biases in AI systems.
- **Transparency.** Companies must be transparent about their use of AI and inform third parties, such as their customers, about how AI is used and the purposes for its application. Specifically, companies that use AI for automated individual decision-making must provide individuals with relevant information in a concise, transparent, intelligible, and easily accessible form regarding the procedure and principles applied to the use of personal data to obtain a specific result (see Chapter 2 for more details).
- **Individuals' rights.** Companies using AI should ensure they respect individuals' rights regarding the processing of their personal data for AI purposes. This involves enabling individuals to access their personal data, correct any inaccuracies, object to the processing, or have their data erased under applicable legal conditions (see Chapter 3 for more details).
- **Security.** Companies using AI must ensure that AI systems are safe before use (see above, *Controller-Processor Agreement*). They must also have processes to notify relevant authorities and affected individuals when necessary (see Chapters 3 and 4 for more details). For instance, if chat logs of an AI chatbot used for customer support containing personal data become publicly accessible due to a misconfiguration or an attack, the company needs to notify both the competent authority and the individuals concerned.

Contact Our Teams

For any additional information on AI or data-related issues under EU law, please contact our teams in Brussels, Frankfurt and London.

BRUSSELS

Anne Vallery
Partner-in-Charge
anne.vallery@wilmerhale.com

Frédéric Louis
Partner
frederic.louis@wilmerhale.com

Itsiq Benizri
Counsel
itsiq.benizri@wilmerhale.com

FRANKFURT

Dr. Martin Braun
Partner
martin.braun@wilmerhale.com

Prof. Dr. Hans-Georg Kamann
Partner
hans-georg.kamann@wilmerhale.com

LONDON

Cormac O'Daly
Partner
cormac.o'daly@wilmerhale.com

Connect with us

wilmerhale.com