AI Act and GDPR: managing the world of data in the world of privacy

Article published with INPLP.

Contrary to some persisting beliefs that the AI Act and GDPR are inherently incompatible, GDPR may in fact be interpreted in a way that concords with the purposes of the AI Act. Processing personal data through an artificial intelligence (AI) system's algorithm triggers the application of GDPR.

I systems are designed to operate with a certain level of autonomy, that is, without human involvement (Recital (6) AI Act).

They infer how to achieve a given set of objectives without being explicitly programmed to attain it, thanks to machine learning, logic-based approaches and knowledge-based methods (Article 3 AI Act).

An AI system requires the training of a computational model to perform specific tasks or to make predictions based on data that has been collected, cleaned, normalised, extracted and validated.

Some AI systems may be trained with data to:

- play a board game,

- to drive vehicles,

- to execute simple voice commands, and

- to generate text-based content.

To that end, AI systems, particularly those rooted in deep learning, rely on vast amounts of data to efficiently:

- identify patterns,

- develop probabilistic models, and

- deliver accurate results.

Data used to train AI systems or provided to them as input often includes personal data, including sensitive personal data, which presents significant privacy concerns and must therefore be compliant with GDPR.

THE QUESTION WE MIGHT ASK IS WHETHER THE AI ACT AND GDPR ARE COMPATIBLE?

The current version of the AI Act explicitly provides that GDPR principles apply to training, validation, and testing datasets of AI systems (Recital (44a) AI Act). Moreover, it underscores that the AI Act in no way affects the obligations of AI providers/users as data controllers or processors (Recital (58a) AI Act). But it is a major challenge for AI system providers to comply with GDPR principles. We will mainly focus on the principle of lawfulness as an exercise to understand whether it is possible to comply with one key GDPR principle.

The principle of lawfulness essentially provides that personal data must be processed in a lawful, fair, and transparent manner (Art. 5.1(a) GDPR). Prior to undertaking data processing, any person or organisation (company, non-profit organisation, foundation, etc.) must therefore identify a legal basis for it from the six bases provided by the GDPR (Art. 6.1 GDPR):

- consent,

- performance of a contract,

- legal obligation,

- vital interests,

- public task, and

- legitimate interests.

Establishing a legal basis for processing personal data can give rise to a complex conundrum within the realm of AI.

(i) Processing based on ‘consent’ (Art. 6.1(a) GDPR): this is only an appropriate lawful basis if the data subject is genuinely offered control and a choice with regard to accepting or declining the terms offered, without any detriment.

The individual’s consent to the processing of their personal data is naturally a highly valued legal basis because it reflects the values of a democratic society. It may indeed seem reasonable to ask the data subject whether they wish to have their personal data processed by another entity.

However, within the context of AI systems:

- ensuring ‘informed consent’ may not always be viable, especially with regards to complex machine learning algorithms whose results are often generated through processes that are not yet fully understood (e.g., the “black box” problem).

- obtaining ‘unambiguous consent’ from every single data subject may also be an impractical approach because of the large datasets, of various categories, originating from countless sources. This issue is exacerbated when data has not been directly obtained from data subjects themselves but has been scraped from websites or obtained through an intermediary instead – including data pools.

- granting data subjects the ‘right to withdraw’: Managing and implementing this right can pose technical difficulties due to the vast quantities of words, images, or sounds. This complexity extends to the fact that each word is processed in a tokenized form within an AI system, either as a single element or broken into multiple separated elements, and only becomes correlated with others when vectorized in a pre-trained machine learning model.

(ii) The ‘performance of a contract’ (Art. 6.1(b) GDPR): this lawful basis for the processing of personal data could be a solution, particularly whenever an existing contract governs the relationship between the provider of an AI system and the end user.

(iii) A legal obligation to which the controller is subject (Art. 6.1(c) GDPR): this is, to the best of our knowledge, not a valid legal basis as there is no law yet requiring the data controller to use AI to meet its legal obligation.

(iv) Alternative legal basis may not always be valid. These specific legal bases require establishing the necessity of the processing for a specific aim that often remains unmet with AI-related processing (Art. 6.1(c), (d) & (e) GDPR).

A processing based on ‘public task’ or to protect ‘vital interests’ may not be valid, particularly when using AI systems for commercial purposes, for instance.

(v) More generally, providers of AI systems could potentially rely on the ‘legitimate interests’ legal basis as a last resort (Art. 6.1(f) GDPR).

This basis would require a careful assessment and may only apply when no other basis finds application (Recital (47) GDPR).

Furthermore, opting for this legal basis demands a careful balance between the interests of providers of AI systems and the data subject’s fundamental rights and freedoms (Art. 6.1(f) GDPR; Recital (47) GDPR).

In any event, an AI system cannot be trained and be designed to operate based on data collected illegally.

Personal data collected in violation of GDPR rules, or any data type that has been indiscriminately scraped from websites, in contravention of a website’s terms of use or of protected databases (see Directive 96/9/EC), could potentially lead to the AI system being banned.

The road to determining an appropriate legal basis for AI solution may be winding and tricky but it does not seem impossible - which is good news.

Other questions are also of great importance when analysing GDPR principles for AI systems. For example, the processing of personal data for purposes other than those for which the personal data have been collected is essential to AI (in particular, for statistical analytics), due to the vast and diverse repositories of data and methodologies deployed to discover correlations and/or potential causal relationships.

In the context of AI, compliance with GDPR principles is crucial to ensure ethical and legal

personal data processing. The misuse or illegal collection of data may lead to serious consequences, including a potential prohibition of AI systems. It is also essential to balance the benefits of AI with the fundamental rights of data subjects to foster responsible development of AI and maintain the privacy and freedoms of individuals in the digital era.

Experts :

Michel MOLITOR

MANAGING PARTNER

Avocat à la Cour, Member of the Luxembourg Bar, 1985

Virginie LIEBERMANN

COUNSEL

Avocat à la Cour, Member of the Luxembourg Bar, 2008

Ruben MENDES

SENIOR ASSOCIATE

Avocat à la Cour, Member of the Luxembourg Bar, 2016