Processing personal data using AI Systems Part 1

In the first in a series of articles, Tim Hickman, Partner, and Aishwarya Jha, Associate, at White & Case LLP, discuss the impact of the EU's AI Act on data protection, starting with the concept of an 'AI System'

The next article in this series will explore the very broad extraterritorial nature of the AI Act.

The EU's Artificial Intelligence Act (the 'AI Act') is nearing final form, with the compromise text (the 'draft text'), the product of trialogue negotiations between EU legislative bodies, having been leaked in January 2024. The overlap between the AI Act and the GDPR/UK GDPR remains somewhat unclear. On the one hand, Article 2(5b) of the AI Act states that the AI Act 'shall not affect' the GDPR. On the other hand, it appears that any organisation using AI Systems — a term explained below — to process personal data will potentially be subject to the overlapping compliance obligations of both the GDPR/UK GDPR and the AI Act.

In order to address the practical consequences of this overlap, it is essential for organisations to understand whether their data processing activities involve the use of any AI Systems within the meaning of the AI Act. This is because the question of whether the AI Act applies will depend, to a large extent, upon the question of whether an AI System is in use. However, as will become clear, defining AI Systems can be beguilingly complicated.

Defining AI Systems

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

It is notable that almost none of the wording from the European Commission's original proposal survived into the draft text. The final definition is similar (but not identical) to the definition proposed by the Organisation for Economic Co-operation and Development ('OECD') in November 2023. The degree to which this definition has been revised during the trilogue process illustrates the difficulties that arise when attempting to pin down the concept of an AI System in technology-neutral terms.

As things currently stand, organisations have little choice but to attempt to apply the definition in the hopes of understanding whether their data processing activities involve any AI Systems within the meaning of the AI Act. To support that assessment, we have considered each of the key phrases in the definition in the sections that follow. As a preliminary, organisations should note that the definition of AI Systems in the AI Act does not depend upon whether the relevant system (or its developer or deployer) is based in the EU.

'A machine-based system'

The first element of the AI Act's definition of an AI System is that it is 'machine-based'. Recital 6 to the AI Act explains that this 'refers to the fact that AI systems run on machines', which is helpful as far as it goes, but leaves open the question of whether a system must run exclusively on machines in order to fall within the definition. For example, many AI models rely extensively on large networks of human reviewers to continually categorise and tag content for inclusion in the relevant AI model. Is a system still 'machine-based' if it cannot function as intended without constant input from humans? As with many of the aspects of the AI Act, at present it is difficult to answer this question with any great degree of certainty.

'...designed to operate with varying levels of autonomy'

Upon a plain reading of the draft text, only AI systems 'designed' to operate with varying levels of autonomy are in-scope. A literal interpretation suggests that AI systems that were not 'designed' to operate with a degree of autonomy are out-of-scope, even if they later develop a degree of autonomy in practice (for example, due to unauthorised or inadvertent human intervention). It is also uncertain what will happen if a system is 'designed' to operate with some level of autonomy, but fails to achieve any autonomy. Is the design element, in isolation, sufficient to satisfy the definition, even if that design is never realised?

Taking into account the overall objectives set out in Article 1 of the AI Act, we anticipate that EU courts and regulators will apply some flexibility to the term 'designed'.

The meaning of the expression 'varying levels' of autonomy is not entirely clear. Arguably, 'varying' could mean that in-scope AI systems must have fluctuating levels of autonomy to fall within the definition, but this seems unlikely to have been the intended meaning. A more likely interpretation is that a system with any level of autonomy (from zero autonomy to complete autonomy) is potentially in-scope.

According to Recital 6, 'autonomy' means 'independence of actions from human involvement and [having] capabilities to operate without human intervention'. It therefore appears that a system will be treated as having some level of autonomy if it has the capacity to perform relevant tasks without human intervention. However, when viewed in conjunction with the terms 'designed' and 'varying' discussed above, it seems that even systems that do not exhibit any autonomy are not necessarily excluded from the scope of the AI Act's definition of an AI System.

'…that may exhibit adaptiveness after deployment'

Recital 6 clarifies that 'adaptiveness after deployment' refers to a system's 'self-learning capabilities, allowing the system to change while in use'. The OECD provides examples that include recommender systems that adapt to an individual's preferences, and voice recognition systems which adapt to a user's voice. However, the term 'may' appears to be entirely permissive in this context — a system may exhibit adaptiveness, or may not, and could still fall within the AI Act's definition of an AI System.

'…for explicit or implicit objectives'

The expression 'explicit or implicit objectives' appears at first glance to simply mean 'any imaginable objectives'. However, Recital 6 clarifies that the term 'implicit objectives' is meant to capture objectives being achieved by a system which are 'different from the intended purpose' of that system. In other words, by including a reference to 'implicit objectives', the AI Act is clarifying that the definition of AI Systems includes both systems that produce intended outputs and systems that produce unintended outputs.

'…infers, from the input it receives, how to generate outputs'

Recital 6 emphasises that the 'capability to infer' is a key characteristic of an AI System, and explains that this term refers to the process of deriving outputs (which, as noted below, are described in non-exhaustive terms in the definition) from inputs (which can include data from any source). This includes the ability to derive models and/or algorithms from such inputs. Recital 6 further states that the techniques that enable a system to 'infer' in this context includes various forms of machine learning, and 'goes beyond basic data processing, enable learning, reasoning or modelling.' This description of the concept of 'inferring' is very broad, and its limits are not easily identifiable.

'…such as predictions, content, recommendations, or decisions'

This phrase appears to be a non-exhaustive list of examples of the types of outputs that an in-scope AI System might produce. However, it appears that an AI System could produce outputs falling into any of these categories, or none of them, and still potentially fall within the definition of an AI System.

'...that can influence physical or virtual environments'

The meaning of this expression is unclear, and the Recitals do not provide further clarity. On a literal reading, the expression 'can influence' appears to mean that an in-scope AI System must at least be capable of some degree of influence, but does not appear to require that such influence actually arises. In addition, it is not certain whether the expression 'physical or virtual environments' simply means 'any conceivable environment', or whether it involves something more specific.

Conclusion

As set out above, the AI Act's definition of an AI System leaves considerable room for doubt. This is likely to be a major challenge for any organisation that is attempting to figure out whether or not its data processing activities fall within the scope of the AI Act. Unless the final version of the text is materially revised, it is likely that this level of uncertainty will remain until the European Commission, and the relevant EU regulators and courts, provide guidance and/or enforcement decisions that will give clarity to the interpretation of the definition.

To an extent, there are parallels with the way in which several terms used in the GDPR have been clarified by the European Data Protection Board and the Court of Justice of the EU over the last few years. However, it should be noted that the uncertainty surrounding the AI Act's definition of an AI System appears to be an order of magnitude greater than any uncertainty relating to foundational concepts in the GDPR.

In an effort to consider whether any better definitions might exist, we asked an AI system to propose its own definition. We leave you to make your own mind up as to whether you prefer its suggestion: 'An AI System is a marvel of computational ingenuity — a cosmic librarian, sifting through the celestial scrolls of information, weaving narratives, and whispering sagas of insight. It is the oracle at the crossroads of logic and intuition. It dances with data, orchestrating a symphony of predictions, content, and decisions. Its essence lies in deciphering the cryptic language of inputs, unravelling their secrets, and conjuring outputs that ripple through the fabric of reality. It peers into the quantum fog, discerning patterns invisible to mortal eyes.'

Tim Hickman and Aishwarya Jha
White & Case LLP
tim.hickman@whitecase.com
aishwarya.jha@whitecase.com