A deep dive into the EU AI Act

15 April 2024

Key takeaways:

- The aim of the Act is to regulate the development and use of AI systems by providing a horizontal framework of obligations for AI developers and AI users ranging from mere transparency requirements to obtaining a comprehensive CE marking. This framework is based on a risk-categorisation of AI systems (low, limited, high) including a prohibition on certain types of AI that pose unacceptable risk.
- The final version of the Act also addresses “general-purpose AI” (GPAI), or models which can be used for a variety of tasks, with the most powerful GPAI models – referred to as ‘systemic’ – subject to additional requirements.
- Special attention needs to be drawn to the requirement imposed by the Act to GPAI to ‘comply with EU copyright law’ which applies on a territorial and on an extraterritorial basis to “ensure a level playing field”. This is unprecedented with respect to a right that is inherently territorial and a significant step towards trying to establish EU copyright law as a global standard.

Authors: Sophie Goossens

What is the EU AI Act?

In April 2021, the European Commission tabled a proposal for the AI Act (the “**Act**”). In typical Brussels fashion, the Act was then under discussion for over two years until December 2023 when a political agreement was finally reached. The agreement was approved by the EU Parliament in March 2024 under the accelerated procedure; it is now being translated in the official languages of the EU and expected to be formally adopted during the plenary session of 22 April.

The aim of the Act is to regulate the development and use of AI systems by providing a horizontal framework of obligations for AI developers and AI users ranging from mere transparency requirements to obtaining a comprehensive CE marking. This framework is based on a risk-categorisation of AI systems (low, limited, high) including a prohibition on certain types of AI that pose unacceptable risk. The final version of the Act also addresses general-purpose AI (GPAI), a term that has been preferred to ‘foundation models’ to describe models which can be used for a variety of tasks, with the most powerful models - referred to as 'systemic' - subjected to additional requirements.

Crucially, the proposed Act is a regulation, meaning that it will be directly applicable in the European Union (EU) Member States and will not need to be transposed into national law. With its entry into force, the regulation will then become part of national law and will be enforceable through national courts of each Member State.

When will the Act come into force?

Once adopted, the Act will come into effect two years after its publication (in 2026), with some particular provisions taking effect sooner: prohibitions will apply after six months, while most rules regarding general-purpose AI, governance, notified bodies, and sanctions will apply after twelve months. For product components covered by sectoral legislation, such as toys, cars or medical devices, and for GPAI model already on the market at the date of the Act, the implementation deadline will be thirty-six months. See our AI timeline on reedsmith.com.

Who does the Act apply to?

Ratione personae

The Act will broadly apply to 'providers' and 'deployers' of AI systems. However, the text envisages that 'importers' and 'distributors' of AI systems, 'product manufacturers' of products with AI inside as well as 'authorised representatives' of providers that are not established in the EU will also be covered. Individuals using AI systems in the course of personal, non-professional activities have no obligation under the AI Act.

- **Providers** are entities located anywhere in the world, who develop AI systems, or a GPAI model, with a view to placing them on the market or putting them into service in the EU.
- **Deployers** (an expression which has been preferred to that of 'user') are natural or legal persons located in the EU using AI under their authority, in the context of a professional activity, externally or internally. Note that a deployer can become a provider upon affixing its own trademark to an AI system, modifying an AI system or modifying the intended purpose of an AI system.
- In addition, the Act envisages that deployers and providers outside the EU may also be covered **where the output produced by an AI system is used in the EU**. This is chiefly to avoid circumventing the Act by merely subcontracting AI tasks or AI data processing to entities based outside of the EU.

Ratione materiae

The Act applies to all AI systems, defined as "machine-based system […] 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", with a few notable exceptions. In particular, the Act does not apply to:

- AI systems developed or used exclusively for military purposes (Military AI);
- AI systems and GPAI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development (Research AI);
- Low risk AI systems released under free and open source licences (low risk open-source AI) and systemic GPAI models released under free and open source licences (regular open-source GPAI models), save that all GPAI models, open source or not, will be subject to the 'Copyright Requirements' (defined below);
- AI systems or models undergoing research, testing and development activities prior to being placed on the market or put into service (Pre-release Testing).
- AI systems used by public authorities in a third country or by international organisations in the framework of international cooperation or agreements for law enforcement and judicial cooperation with the Union or with one or more Member States, under the condition that this third country or

international organisations provides adequate safeguards with respect to the protection of fundamental rights and freedoms of individuals (International Cooperation AI).

The definition of an AI system was updated a number of times to align it more closely with the work of international organisations working on AI, including the OECD, and will be the subject of guidelines from the Commission, who already specified that it is not intended to cover simpler traditional software systems or programming approaches based on rules defined solely by natural persons to automatically execute operations.

What is the risk-based approach followed by the Act?

The Act provides for rules based largely on risk, so the higher the perceived risk, the stricter the rules. The Act classifies AI systems into four risk categories: unacceptable risk, high risk, limited risks, and low or minimal risk.

- AI systems with an **unacceptable** level risk are prohibited. This includes the most intrusive uses of AI systems, for instance, emotion recognition systems in the workplace or education, systems for assessing the risk of committing a criminal offence, social scoring or biometric categorisation systems.
- AI systems with a **high risk** will be subject to the strictest requirements under the Act which include the need to apply for a CE marking. There are three ways in which the legislation provides for AI systems to be considered 'high-risk': (i) when the AI system is itself a certain type of product; (ii) when the AI system is a safety component of a certain type of product; (iii) when the AI system meets the description of listed 'high-risk' AI systems.
    - **Products** containing AI including medical devices, industrial machinery, toys, aircraft or cars, among other examples will be deemed high risk, where those products are already subject to certain EU regulation, or when they are required to undergo a third party conformity assessment before it is placed on the market or put into service in the EU.
    - Similarly, AI enabled **safety components of products** (e.g., medical devices, industrial machinery, toys, aircraft, or cars) or of equipment (e.g., rail infrastructure, lifts, or appliances burning gaseous fuels, etc.) will be deemed high risk where those products are already subject to certain EU regulation, or when they are required to undergo a third party conformity assessment before being placed on the market or put into service in the EU.
    - Finally, **AI systems used** in the following **eight categories** may be deemed high risk further to an assessment of the risk and harm they pose: biometrics, critical infrastructure, education, employment, access to essential services (both public and private), law enforcement, immigration, administration of justice and democratic processes.

  As above, if an AI system falls in one of the categories, it may still be excluded from the high risk category if it "does not pose a significant risk of harm, to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making". This is the case if the AI system meets at least one of four criteria, based on the objective of the AI system used including if the system: (i) performs a narrow procedural task; (ii) improves the result of a previously completed human activity; (iii) detects decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human

assessment, without proper human review; or (iv) performs a preparatory task to an assessment relevant for the purpose of the use cases that would otherwise be high-risk.

- AI systems with **limited risk** are all AI systems which are not high risk but that may still display manipulation or deceit – this includes in particular chatbots, deepfakes systems or synthetic content generation systems. These systems will not be subjected to the CE marking procedure but will nonetheless need to meet specific transparency requirements to ensure humans are appropriately informed of their capabilities. The requirements, set out at article 50, are as follows:
    - AI systems interacting directly with natural persons must contain visible information about it, unless this is obvious;
    - AI systems (incl. GPAI models) generating synthetic content must make sure their outputs are marked in a machine-readable format and detectable as AI-generated. However the requirement does not apply to assistive function for standard editing, or where input data was not substantially altered;
    - Emotion recognition or biometric systems must bear visible information to the natural persons and process their personal data in accordance with the GDPR;
    - Deep fake systems must disclose that their outputted content has been AI generated or manipulated. However for evidently artistic, creative, satirical, fictional or analogous work or programme, the requirement is limited to disclosing the existence of the deepfake or synthetic content “in an appropriate manner that does not hamper the display or enjoyment of the work”;
    - News outlets (informing the public “on matters of public interest”) using text-based AI systems must disclose that the text has been AI generated or manipulated. However the requirement does not apply where the content has, cumulatively, undergone a process of human review or editorial control and where the new outlet assumes editorial responsibility for the content.
- AI systems with a **low risk** are all other AI systems, including as AI-enabled recommender systems or spam filters; they are largely unregulated.

What are the obligations for high-risk AI systems?

High-risk AI systems must comply with several mandatory requirements before the system can be placed on the market or put into service, or before its output can be used in the EU. High-risk systems are subject to a conformity assessment – and a CE marking process – that is intended to certify that the system in question meets these requirements. The requirements are onerous and include in particular: the creation of risk management and data governance procedures, drawing up technical documentation, ensuring the system has logging capabilities, demonstrating human oversight, accuracy, robustness and the cybersecurity resilience of the system.

What are General Purpose AI models?

The release of ChatGPT at the end of 2022 opened the world’s eyes to the phenomenal versatility of LLMs and prompted the Members of the EU Parliament (MEPs) to add a new definition to the Act to better capture this family of systems and impose additional requirements to it. This was done by adding the concept of General Purpose AI models and systems (an expression that has been preferred to that of ‘foundation models and systems’) to the text, defined as follows:

- **general purpose AI model** is an “AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications”, and
- **general purpose AI system** is an AI system “based on a general purpose AI model, that has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems.”

GPAI models will fall into two categories: ‘systemic’ or not, an assessment that will be made either based on a quantitative threshold of the cumulative amount of compute used for its training, or on an individual designation decision of the Commission. Systemic GPAI models are subject to additional requirements including performing model evaluation, making risk assessments, taking risk mitigation measures, ensuring an adequate level of cybersecurity protection, and reporting serious incidents to the AI Office and national authorities.

The creation of the GPAI status has enabled MEPs to address the concerns of two important categories of EU stakeholders: (i) EU entities using GPAI models in downstream applications, and (ii) copyright rightsholders.

- For the former, the text now obliges providers of GPAI models to draw up and supply to their downstream users the technical documentation of the model. This includes details and elements of high-risk AI systems, enabling users to fulfill their respective requirements, particularly for conformity assessment purposes.
- For the latter, the text combines two powerful requirements (the Copyright Requirements):

(i) an obligation to draw up **a policy to comply with EU copyright law**, including by identifying and complying with the reservation of rights (‘opt out’) permitted under article 4(3) of the Copyright Directive; and

(ii) an obligation to draw and make publicly available **a sufficiently detailed summary about the content used for training of the general-purpose AI model**, according to a template provided by the AI Office. The summary should be “generally comprehensive in its scope instead of technically detailed to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under Union law, for example by listing the main data collections or sets that went into training the model, such as large private or public databases or data archives, and by providing a narrative explanation about other data sources used.”

Extraterritorial application of EU copyright law

Special attention needs to be drawn to the requirement imposed by the Act to ‘comply with EU copyright law’ and in particular to the recital which has been adopted alongside it, declaring the application of EU copyright law on a territorial *and* on an extraterritorial basis, a protectionist move whose impact will be widley felt. The recital, which was bitterly fought over during the final stretch of the negotiation of the Act, has seemingly been adopted to “*ensure a level playing field among providers of general-purpose AI models where no provider should be able to gain a competitive advantage in the Union market by applying lower copyright standards than those provided in the Union.*” This is unprecedented with respect

to a right that is inherently territorial and a significant step towards trying to establish EU copyright law as a global standard. It is unlikely to go unnoticed.

What are the penalties provided for by the Act?

For organisations caught by the Act, strict financial penalties for non-compliance can be imposed. While the highest fine of up to EUR 35 million or 7% of worldwide annual turnover is limited to non-compliance with the AI prohibitions, non-compliance with the transparency requirements or with the GPAI requirements can result in a fine of up to EUR 15 million or 3% of worldwide annual turnover, whichever is higher.

In-depth 2024-078