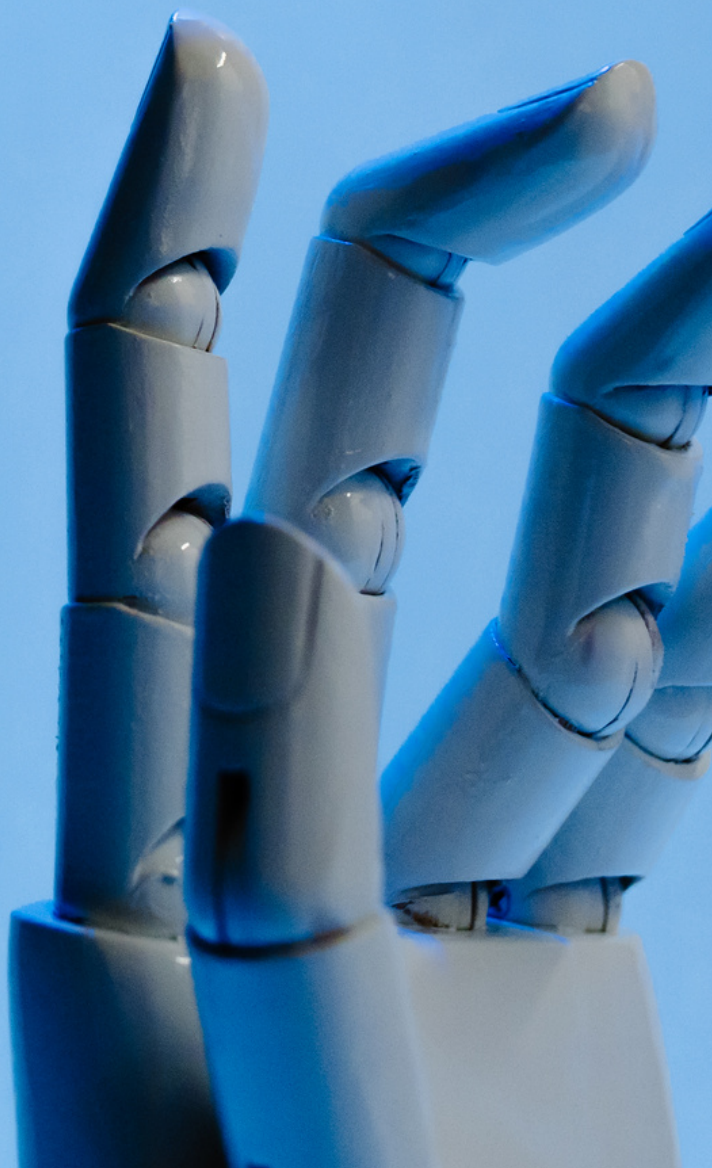
Giulio Coraggio

AI Act Finalized

Here is what has been agreed



DLA PIPER

AI Act Finalized - Here is what has been agreed

The technical experts of the European Union have finalized the **EU AI Act,** the first legislation to regulate the much-discussed artificial intelligence (AI) that **reached its final version**.

After last December's marathon of three-day negotiations by the trialogue composed of the European Commission, Council, and Parliament, the technical experts of the European Union have extensively worked to reach the **now finalized text of the AI Act**. We will have to wait for the official approval by the EU Parliament of the first regulation in the world on artificial intelligence that will probably happen in April 2024, but it appears to be a mere formality.

The definition of artificial intelligence system under the finalized AI Act

There is a new definition of artificial intelligence system that slightly differs from the one provided by the OECD guidelines and is the following:

> *"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"*

The main 3 components of the definition are:

1. both systems where the AI operates with **full autonomy and under human inputs**;
2. systems that may **adjust as a consequence of the information provided** to them and
3. systems that, based on the information received, **learn how to generate outputs**.

The goal of the EU legislator is to adopt a **definition of AI systems that is as broad as possible**. It gives it a broad scope since artificial intelligence could impact every sector, excluding from its application some systems already subject to harmonization legislation as well as those exclusively used for military, defense, or national security purposes, in addition to systems used for scientific research, development, and purely personal non-professional usage.

A further highly discussed exception to the applicability of the AI Act operates in relation to **AI systems exploiting free and open source software** that are not subject to the terms of the regulation unless it is

1. either put on the market or put into service as a **high-risk AI system**; or
2. subject to the **transparency obligations imposed under the AI Act**.

Also, open source components that can benefit from the exemption are those whose parameters, including weights, on model architecture and model usage are made publicly available and are not made available against a price or otherwise monetised.

In any case, the free and open source exemptions do not apply if the AI system is designated as GPAI with systemic risks.

The classification of AI systems

The legal framework for AI is characterized by a dual regime distinguishing between AI systems of limited risk and those with high risk. The definition of risk is "*the combination of the probability of an occurrence of harm and the severity of that harm*," and, based on such definition, **there is a distinction among**

Prohibited AI Systems

These systems include (i) techniques that manipulate individual cognition and behavior, (ii) the random collection of facial recognition data from the Internet or through CCTV, (iii) the use of emotion recognition systems in workplaces and educational settings, (iv) the deployment of social credit scores and (v) the biometric processing for the inference of sensitive personal data like sexual orientation or religious beliefs. **Such AI Systems are just banned**.

High-Risk AI Systems

These systems include:

- AI systems intended to be used as a **safety component of a product**;
- AI systems falling under harmonized legislation listed in **Annex II of the AI Act**, and
- some AI systems used in the **educational sector**,
- in **recruiting and employment processes as well as for credit scoring purposes**, unless there is a fraud detection purpose, and risk assessment and pricing concerning natural persons in the case of life and health insurance and for AI systems used for other purposes listed in Annex III of the AI Act.

Such categories are not rigid, though, since the provider can prove that the specific system is not high-risk because of its peculiarities.

A risk management system shall be established, implemented, documented, and maintained in relation to high-risk AI systems to identify risks and adopt mitigating actions throughout the whole life-cycle of the AI system, also performing tests to understand the operation of the system in real-world conditions. If the high-risk AI system involves the training of models with data, it shall be developed based on training, validation, and testing data sets subject to appropriate data governance and management practices listed in the AI Act. The performance of these activities in relation to high-risk AI systems shall be proven **through technical documentation** to be arranged before the system is placed on the market or put into service and shall be kept up-to-date. Such technical documentation shall at least contain the information in Annex IV of the AI Act.

Still, it represents a **self-assessment, with no reference to a third party assessment.** Such assessment shall also be supported by evidence gained through the automatic recording of events ('logs') throughout the system's lifetime that the system shall technically allow. The high-risk AI systems shall be designed and developed in such a way to

- Ensure that their **operation is sufficiently transparent** to enable deployers to interpret the system's output and use it appropriately, also thanks to instructions that shall accompany the system;
- Allow an **effective oversight by natural persons** during the period in which the AI system is in use, also with appropriate human-machine interface tools;
- Achieve an **appropriate level of accuracy, robustness, and cybersecurity**, and perform consistently in those respects throughout their lifecycle;
- Be able to provide **individuals**, in case of automated decisions, an **explanation of the decision-making procedure** and the main elements of the decision taken.

General Purpose AI Systems

These systems "*are 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*" and are meant to **carry a systemic risk** (i.e., a risk with adverse effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain), when either are identified

as such by the EU Commission or when the cumulative amount of compute used for its training measured in floating point operations (FLOPs) is greater than 10^25.
GPAI systems generating a systemic risk shall be notified to the EU Commission. Besides, providers of GPAI systems shall: Draw up and keep up-to-date the technical documentation of the model and make available information and documentation to providers of AI systems who intend to integrate the general-purpose AI model in their AI system. Such technical documentation shall be drafted listing, among others, the:

- modalities followed to develop the system, the activities performed, and the estimated energy consumption, and – in case of GPAI systems carrying a potential systemic risk the evaluation strategies that have been followed and the adversarial testing (e.g., red teaming) performed;
- Put in place a **policy to respect Union copyright law** and draw up and make publicly available a sufficiently detailed summary of the content used for training of the general-purpose AI model, according to a template provided by the AI Office; and
- In case of **GPAI with systemic risk**, (i) perform **model evaluation**, (ii) assess and **mitigate possible systemic risks** at the Union level, (iii) keep **track of, document, and report** relevant information about serious incidents and possible corrective measures to address them and (iv) ensure an **adequate level of cybersecurity protection**.

Basic AI Systems

All AI systems, regardless of the level of risk, are subject to minimum obligations.
They are subject to basic transparency obligations to ensure a minimal level of clarity and understanding across the board, informing individuals that they are interacting with an AI system.

Which are the entities obliged under the AI Act?

The finalized EU AI Act has a transnational effect as it applies to:

1. Providers placing on the market or putting into service AI systems or placing on the market general-purpose AI models in the Union, **irrespective of whether those providers are established or who are located within the Union or in a third country**;
2. Deployers of AI systems that **have their place of establishment or who are located within the Union**;
3. Providers and deployers of AI systems that **have their place of establishment or who are located in a third country, where the output produced by the system is used in the Union**;
4. Importers and distributors of AI systems;
5. Product manufacturers placing on the market or putting into service an AI system together with their product and under their own name or trademark;
6. Authorized representatives of providers, which are not established in the Union; and
7. Affected persons that are located in the European Union.

The **most relevant definitions of the categories of entities** referred above are those of:

A) The provider

The "*provider*" that is a "*natural or legal person, public authority, agency or other body that **develops** an AI system or a general purpose AI model or that **has** an AI system or a general purpose AI model **developed and places them on the market** or **puts the system into service under its own name or trademark**, whether for payment or free of charge*".

The provider is the entity subject to the **most relevant obligations under the AI Act** since, among others, it needs to ensure that their **high-risk AI systems are compliant** with the requirements of the Act, also

- having in place a quality management system,
- keeping the documentation to be able to prove compliance with the AI system,
- adopting the corrective actions if the AI system is not compliant with the Act,
- drawing up a written machine readable, physical or electronically signed EU declaration of conformity for each high-risk AI system;
- registering high-risk AI systems with the EU database; and
- complying with the obligations applicable to GPAI systems.

B) The deployer

> The *"deployer"* that is *"any natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity"* which includes any company that receives an AI system by a supplier to run its operations.

To **avoid any lack of responsibility of deployers**, the AI Act provides that they shall not only:

- take **appropriate technical and organisational measures** to ensure they use systems in accordance with the instructions of use accompanying the systems;
- assign **human oversight to natural persons** who have the necessary competence, training and authority, as well as the necessary support;
- **monitor the operation of the AI system**; and
- **comply with information obligations** before the AI system is put in operation and perform a DPIA when required by the type of processing of personal data.

Also, deployers of high risk AI systems that, among others, are used for credit scoring as well as for risk assessment and pricing in relation to natural persons in the case of life and health insurance shall perform the **fundamental rights impact assessment (FRIA)** that shall be notified to the market surveillance authority with the assessment results, submitting the relevant filled template referred to in the Act.

Central and Local AI Governance

In terms of governance and compliance, the AI Act establishes a **European AI Office to monitor the most complex AI models**. It provides for the creation of a scientific panel and an advisory forum to integrate the perspectives of the different stakeholders. This ensures that regulation is always informed and up-to-date with respect to developments in the field.

But a topic of considerable discussion will be about what powers in practice are given to local AI authorities and which entities will be appointed as national authorities. As happened with GDPR, local authorities will not want to give up their powers. **The AI Office should reduce the risk of inconsistent approaches across the EU among local authorities**, but political friction between different local authorities cannot be ruled out.

The potential sanctions based on turnover

The finalized AI Act also, of course, establishes a system of penalties that, as has been the case with several recent European regulations, is based on companies' global turnover or a predetermined amount, whichever is higher.

Infringements on prohibited practices or non-compliance relate to requirements on banned AI applications	up to **35 million euros** or **7% of global annual turnover** of the preceding financial year
Non-compliance with any of the other requirements or obligations, including infringement of rules on GPAI	**15 million euros** or **3% of global annual turnover** of the preceding financial year
Supply of incorrect, incomplete or misleading information to notified bodies and national competent authorities in reply to a request	**7.5 million euros** or **1.5% of global annual turnover** of the preceding financial year

Exceptions are made for smaller businesses, with limited penalties for SMEs and startups. Thus, even on penalties, a balance has been struck between the need to regulate AI and the goal of not restricting the development of this technology in the EU. For the same reason, the so-called "sandboxing" solutions are provided where solutions can be tested while benefiting from a special regime.

The timeline of the AI Act

6 months of transition period

April 2024
Publication of the AI Act final text

~ October 2024
Introduction of AI bans

~ April 2025
Provisions on GPAI systems enter into force

~ April 2026
Remaining provisions enter into force, save for the below

~ April 2027
Provisions applicable to AI systems used in sectors where there are already EU harmonized rules

The applicability date of the finalized AI Act will follow a precise timeline, with a **transition period of six months** for the introduction of bans, **one year for GPAI systems**, and **two years for the remaining provisions**, except for provisions applicable to devices that are **already regulated by other EU harmonization regulations** for which the time **limit is 36 months**, such as the pharma and the medical devices sector.

The final text will be published in the Official Journal of the European Union by April 2024, at which time the above deadlines will begin to run. However, there is no doubt that regardless of the length of the transition period, no company will be willing to adopt AI solutions that do not comply with the AI Act that would force it to divest from the technology anytime soon.

We built an innovative solution to support businesses in ensuring artificial intelligence compliance in a cost-effective and efficient manner, you can read more HERE and reach out to us to know more information. Also, you can watch HERE a webinar that I ran with my DLA Piper colleagues on the topic.