February 05, 2024

EU Artificial Intelligence Act — Final Form Legislation Endorsed by Member States

Authors: Huw Beverley-Smith, Charlotte H N Perowne

At a Glance

- The EU Artificial Intelligence Act will provide the following:
  - A broad definition of AI that applies to many different entities, including providers, deployers, importers and distributors of AI systems, and will have a wide extra-territorial scope
  - A cross-sectoral, risk-based classification system with an outright prohibition on certain AI practices deemed to impose unacceptable risk
  - New obligations largely targeting AI systems deemed “high-risk” and obligations on providers of general purpose AI systems, including generative AI systems like ChatGPT
  - Significant fines similar to the GDPR of up to 7% of annual global turnover for certain offences

The long-awaited proposed AI Act, once enacted, will be a comprehensive cross-sectoral regulatory framework for artificial intelligence (AI). Its aim is to regulate the development and use of AI by providing a framework of obligations for parties involved across the entire AI supply chain. As with the General Data Protection Regulation (GDPR), the EU is seeking, through its first-mover advantage, to set the new global standard for AI regulation.

Affected Entities

The AI Act is designed to cover the whole AI supply chain. There are a broad range of “operators”, defined as the providers, product manufacturers, deployers, authorised representatives, importers or distributors. The AI Act has wide extra-territorial scope, therefore, a number of US businesses will be within scope, depending on their exact role in the supply chain.

Scope and Definition of AI

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.”

The EU has adopted the definition promulgated by the Organisation for Economic Cooperation and Development (OECD). Unlike the Commission’s original proposal, where an AI system was limited to software acting for human-defined objectives, this now also seeks to encompass the metaverses through the explicit inclusion of virtual environments.

Risk Classification

The AI Act classifies AI systems into several risk categories, with different degrees of regulation applying to each one. Most AI systems will fall into the lower risk categories, although there will be significant procedural and transparency obligations relating to their use, which businesses throughout the supply chain will have to understand.

1. Prohibited — AI Systems Posing Unacceptable Risks

While some of these are focused on the activities of governments and public authorities, many will apply more broadly. The prohibited AI include:

- Use of subliminal techniques to modify or distort an individual’s behaviour (for example, using subvisual cues which cannot be detected by the human eye in sales or advertising)
- Targeting and exploiting the vulnerabilities of specific groups (for example toys using voice assistants which encourage dangerous behaviour)
- Biometric categorisation systems (except for labelling or filtering of lawfully acquired biometric datasets, or for law enforcement purposes where biometric datasets have been lawfully acquired)
- Social scoring (for example, through algorithms determining a person’s ability to access social benefits,

services or opportunities)

- Real-time remote biometric identification systems in publicly accessible spaces for law enforcement (subject to specific exceptions such as the detection of serious crimes)
- Predictive policing
- Facial recognition databases compiled through untargeted scraping of the internet or CCTV footage
- Emotion recognition systems in educational or workplace settings (except where this is intended for medical or safety reasons)

2. High-Risk Systems

The bulk of the obligations under the AI Act apply to systems deemed “high-risk” (estimated by the European Commission to be roughly 5-15% of AI systems). These are divided into two categories (1) safety components or products that are already subject to EU safety legislation and which require third-party conformity assessments under such laws, and (2) standalone AI systems expressly designated by the European Commission as “high-risk”, including AI systems used:

- To determine access to or outcomes in education and training
- In employment for recruitment or selection, promotion, or termination
- To determine access to essential private and public services, like health care or credit (although there are exceptions for fraud detection)
- In managing critical infrastructure
- For non-banned biometric identification systems, including systems inferring protected characteristics, emotion recognition systems and remote biometric identification systems (excluding biometric verification of individuals)
- In managing migration, asylum and border controls
- In the administration of justice and democratic processes (provided that final rulings are made by humans)

Standalone AI systems will not be deemed high-risk if they do 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. Prior to placing on the market an AI system, which falls under Annex III but is deemed not to be high-risk, the provider must document its assessment of the risk, as well as registering the system in the EU database. Such assessment should take into account both the severity of the possible harm and its probability of occurrence. An AI system may be found not to be high-risk if the AI system meets certain criteria, for example, if it is only intended to perform a narrow procedural task or improve the result of a previously completed human activity. However, an AI system will ***always*** be considered high-risk if the AI system performs profiling of natural persons.

AI systems which do not fall into the ‘prohibited’ or ‘high-risk’ categories are permitted, but with transparency obligations — for example ensuring that users know that they are interacting with an AI system in a chatbot.

General Purpose AI

The AI Act has had a relatively long legislative process and has been overtaken by technology developments, particularly the recent rise of generative AI systems, such as ChatGPT. One of the key issues in the final stages of negotiations was the extent to which providers of general-purpose AI (GPAI) models, on which a number of downstream applications are based, should be subject to specific rules.

The final draft sets out specific requirements of GPAI models, including classification of models with systemic risk, procedural requirements, and obligations for providers of both regular GPAI models and those classified with systemic risk.

A GPAI model will be classified as a GPAI model with systemic risk if it meets any of the following criteria: (1) it has high impact capabilities evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks; or (2) based on a decision of the European Commission, including following an alert by a scientific panel that a GPAI model has equivalent capabilities or impact equivalent to those of point (1).

High impact capabilities (under point 1) will be assumed for high-powered AI, where the cumulative amount of computation used for training a model measured in floating point operations (FLOPs) is greater than 10^25.

Providers of GPAI models must:

- Draw up and keep up-to-date technical documentation of the model (including training and testing)
- Provide information to downstream providers of AI systems who intend to integrate the GPAI model into their AI system, containing sufficient information to enable providers of AI systems to understand the GPAI model and comply with their own obligations
- Put in place a policy to respect copyright law (especially with regard to reservation of rights)
- Draw up and make publicly available a sufficiently detailed summary of the content used for training the GPAI

model (according to a template to be provided by the AI Office)

- Cooperate with the European Commission and national competent authorities

Providers of GPAIs with systemic risk will also be required to:

- Perform model evaluation in accordance with standardised protocols and tools, including conducting and documenting adversarial testing of the model to identify and mitigate systemic risk
- Assess and mitigate possible systemic risks, including their sources, that may stem from the development, placing on the market or use of GPAI models with systemic risk
- Report serious incidents and possible corrective measures to address them to the AI Office
- Ensure an adequate level of cybersecurity protection

Except for GPAI models with systemic risks (which must comply with all of the above), AI models that are made available to the public under free and open licences meeting certain requirements are not required to provide the technical documentation and information set out in the first two points above.

Timeline

The final text of the AI Act is expected to be formally adopted in the of summer 2024. However, the leaked final drafts give a sufficient indication of the ultimate requirements for businesses to start planning and integrating the AI Act’s requirements in their product development and procurement processes.

Assuming that the text is adopted in June 2024, the AI Act would be effective 24 months after entry into force (i.e., June/July 2026). However, some provisions, such prohibitions on unacceptable-risk AI systems, will apply 6 months after entry into force (i.e., late 2024). Similarly, regulation of GPAI systems will apply after 12 months (or 24 months if they are already on the market).

We continue to monitor developments and will provide updates on the final text once published, including any last-second amendments.

The material contained in this communication is informational, general in nature and does not constitute legal advice. The material contained in this communication should not be relied upon or used without consulting a lawyer to consider your specific circumstances. This communication was published on the date specified and may not include any changes in the topics, laws, rules or regulations covered. Receipt of this communication does not establish an attorney-client relationship. In some jurisdictions, this communication may be considered attorney advertising.

MEET THE AUTHORS

Huw Beverley-Smith

Partner

+44 (0) 20 7450 4551
London
huw.beverley-smith@faegredrinker.com

Charlotte H N Perowne

Associate

+44 (0) 20 7450 4532
London
charlotte.perowne@faegredrinker.com

Related Legal Services

Government & Regulatory

International

Labor & Employment

Education

Privacy, Cybersecurity, Data Ethics & Strategy

Customs & International Trade

International Transactions