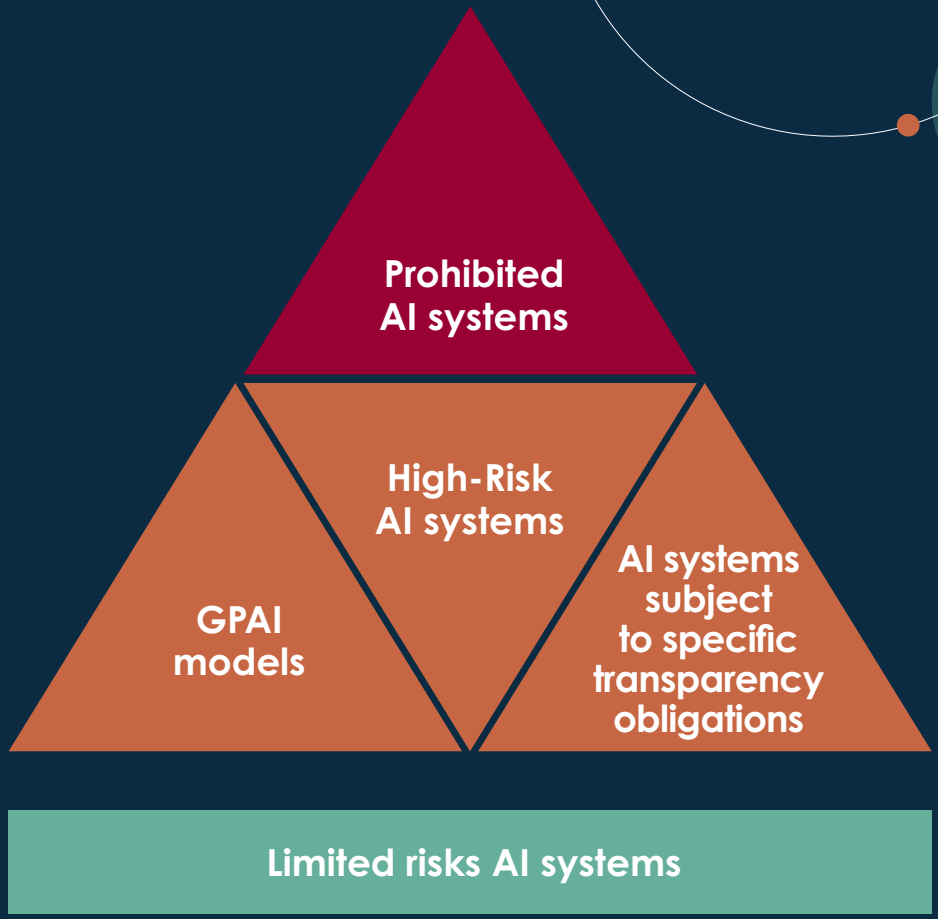
Classification of AI systems and GPAI Models

The EU AI Act classifies AI systems and General-purpose AI (GPAI) models on the basis of the level of risk they pose and the purpose they have. In a nutshell:

- Some AI Systems will be **prohibited** in the EU;
- Some AI Systems will be classified as **high-risk** and will be subject to stringent pre-market and post-market obligations;
- Some AI Systems will be **subject to specific transparency obligations**; and
- Specific rules will apply to **GPAI Models**.



Notes:

- The main purpose of the EU AI Act is to ensure that safe and trustworthy AI systems are used in the EU. Accordingly, EU legislators have decided to adopt a risk-based approach, with the concept of risk being defined as "*the combination of the probability of an occurrence of harm and the severity of that harm*".
- Such a classification is **not mutually exclusive**, e.g. an AI System can be classified as high-risk and be subject to specific transparency obligations at the same time.
- The assessment of the classification of an AI System/GPAI Model must be **performed on a case-by-case basis, documented, reviewed regularly and be kept up-to-date**.
- The EU AI Act does not include any specific obligation for limited risks AI systems; however Providers and Deployers of **limited risks AI systems** will **still be subject to the AI literacy obligation**, which entails taking measures to ensure their staff and other persons dealing with the operation and use of AI systems have the appropriate skills, knowledge and understanding to allow them to make an informed deployment of AI systems, as well as to be aware of the opportunities, risks, and possible harm that AI system can cause.

Prohibited AI Systems

AI systems that deploy subliminal techniques beyond a person's consciousness / purposefully manipulative/ deceptive techniques

which materially distort a person's / group's behaviour by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause significant harm

AI systems exploiting an individual's / group's vulnerabilities, which materially distort their behaviour in a manner that causes or is likely to cause significant harm

AI systems for the evaluation / classification of individuals / groups based on their social behaviour/ personal or personality characteristics, leading to detrimental or unfavourable treatment in unrelated social contexts or unjustified or disproportionate treatment

AI systems to infer emotions of an individual in the workplace and education institutions

⚠ Except if used for medical or safety reasons

AI systems making risk assessments on the risk of a person to commit a criminal offence, based solely on profiling or on assessment of personality traits and characteristics

⚠ Except if used to support the human assessment of the involvement of a person in a criminal activity

AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet / CCTV footage

AI systems for biometric categorisation that individually categorise individuals based on their biometric data to deduce / infer sensitive data

⚠ Except labelling / filtering of lawfully acquired biometric datasets based on biometric data / categorizing of biometric data in the area of law enforcement

AI systems for real-time remote biometric identification in publicly accessible spaces for the purposes of law enforcement

⚠ Except if strictly necessary for limited law enforcement purposes

AI systems infringing other EU laws

Notes:

- The abovementioned AI systems will be **completely banned from the EU from 2 February 2025**, and it will not be possible to place them / put them into service / use them on the EU market.
- It is important to understand the differences amongst the concepts of:
  - **Biometric categorization system** refers to an AI system for the purpose of **assigning individuals to specific categories on the basis of their biometric data**;
  - **Remote biometric identification system** refers to an AI system for the purpose of **identifying individuals**, without their active involvement, typically at a distance **through the comparison of an individual's biometric data with the biometric data contained in a reference database**;
  - **Biometric verification system** refers to the **automated, one-to-one verification, including authentication**, of **an individual's identity by comparing their biometric data to previously provided biometric data**.

  ➔ Only Biometric categorization systems and Remote biometric identification systems meeting the conditions identified above will be prohibited. Conversely, biometric verification systems do not qualify as "Prohibited AI systems".
- Permitted AI systems for real-time remote biometric identification in publicly accessible spaces for the purposes of law enforcement will be subject to onerous obligations (e.g., prior administrative / judicial authorization; fundamental rights impact assessment; etc.).

High-risk AI Systems

- AI Systems intended to be used as a safety component of a product/which are themselves products (i) covered by below EU legislations (referenced in Annex I); and (ii) subject to a third-party conformity assessment procedure

Annex I Section A

Machinery Regulation	Directive on **safety of toys**	Directive on **recreational craft and personal watercraft**
Directive on **lifts and safety components for lifts**	Directive on **equipment** and **protective systems** intended **for use in** potentially **explosive atmospheres**	Directive on **radio equipment**
Directive on **pressure equipment**	Regulation on **cableway installations**	Regulation on **personal protective equipment**
Regulation on **appliances burning gaseous fuels**	**Medical devices** Regulation	**In vitro diagnostic medical devices** Regulation

Annex I Section B

Regulation on common rules in the field of **civil aviation security**	Regulation on **common rules in the field of civil aviation**
Regulation on the approval and market surveillance of **two- or three-wheel vehicles and quadricycles**	Regulation on the approval and market surveillance of **agricultural and forestry vehicles**
Directive on **marine equipment**	Directive on the **interoperability of the rail system within the EU**
Regulation on the approval and market surveillance of **motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles**	Regulation on type-approval requirements for **motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles**

These High-risk AI systems falling under Annex I Section B are **not subject to the requirements for High-risk AI systems laid down under Chapter III Section 2**, of the EU AI Act. The requirements that they will need to comply with will be **integrated into the product legislation's technical specifications and procedures.**

Steptoe | EU AI Act Decoded

High-risk AI Systems

- AI Systems used in the below areas (referenced in Annex III)

Administration of justice and democratic processes

- to be used by judicial authorities / on their behalf to assist in researching and interpreting facts / law application / alternative dispute resolution
- to be used for influencing the outcome of an election / referendum / voting behavior

Access to and enjoyment of essential private / public services and benefits

- to be used by public authorities / on behalf of public authorities regarding access to essential public assistance benefits and services (e.g., healthcare services)
- to evaluate creditworthiness / establish credit score (except if used to detect financial fraud)
- for risk assessment and pricing in the case of life and health insurance
- to evaluate / classify / dispatch / prioritize emergency calls

Law enforcement

- to be used by / on behalf of law enforcement authorities for risk assessment of an individual becoming the victim of criminal offences
- polygraphs or similar tools
- to evaluate the reliability of evidence in the criminal investigation / prosecution
- for risk assessment of an individual offending / re-offending or to assess personality traits and characteristics / past criminal behavior
- for profiling of individual in the detection / investigation / prosecution of criminal offences

Employment, workers' management and access to self-employment

- for recruitment / selection
- to make decisions affecting the work-related relationship

Migration, asylum and border control management

- to be used by / on behalf of competent public authorities as polygraphs or similar tools
- for risk assessment of an individual intending to enter / who has entered into the EU territory
- for the examination of asylum / visa / residence permit applications and associated complaints
- for the purpose of detecting / recognizing / identifying natural persons (except if used for verification of travel documents)

Critical infrastructure

- safety components in the management and operation of critical digital infrastructure, road traffic, or in the supply of water, gas, heating or electricity

Education and vocational training

- to determine access / admission or to assign natural persons to educational and vocational training institutions at all levels
- to evaluate learning outcomes
- for the purpose of assessing the appropriate level of education for an individual
- for monitoring and detecting prohibited behavior of students during tests

Biometrics

- remote biometric identification systems
- biometric categorization according to sensitive / protected attributes or characteristics
- emotion recognition

Exception

An AI system used in one of these areas **will not be considered as high-risk** where it **does not pose a significant risk of harm to the health / safety / fundamental rights** and where it **meets either of these conditions**:

- It is intended to perform a **narrow procedural task**;
- It is intended to improve **the result of a previously completed human activity**;
- It is intended to **detect decision-making patterns or deviations** from prior decision-making patterns and which are not meant to replace or influence the previously completed human assessment, without proper human review; or
- It is intended to **perform a preparatory task** to an assessment.

Notes

- **Safety component** must be understood as a component of a product / AI system which fulfils a safety function for that product / AI system, or the failure / malfunctioning of which endangers the health and safety of persons / property.
- **For High-risk AI systems intended to be used as a safety component** of a product/which are themselves products covered by legislations referred to in Annex I, it will be important to **first assess the applicability of the relevant legislations**.
- For AI systems covered by Annex I Section B, it will be important to monitor the adoption of the product legislation's technical specifications and procedures.
- The **list of High-risk AI systems referred to in Annex III** may be **updated from time to time** by the European Commission.
- A provider who considers that an **AI system referred to in Annex III is not high-risk must document its assessment**. It will however remain subject to the obligation to register its AI systems in the EU database for high-risk AI systems.
- **AI systems referred to in Annex III** that perform **profiling of individuals** will be **always classified as high-risk.**

AI systems subject to specific transparency obligations

AI systems intended to interact directly with natural persons

Except AI systems authorised by law to detect / prevent / investigate / prosecute criminal offences

AI systems generating synthetic audio, image, video or text content (incl. GPAI)

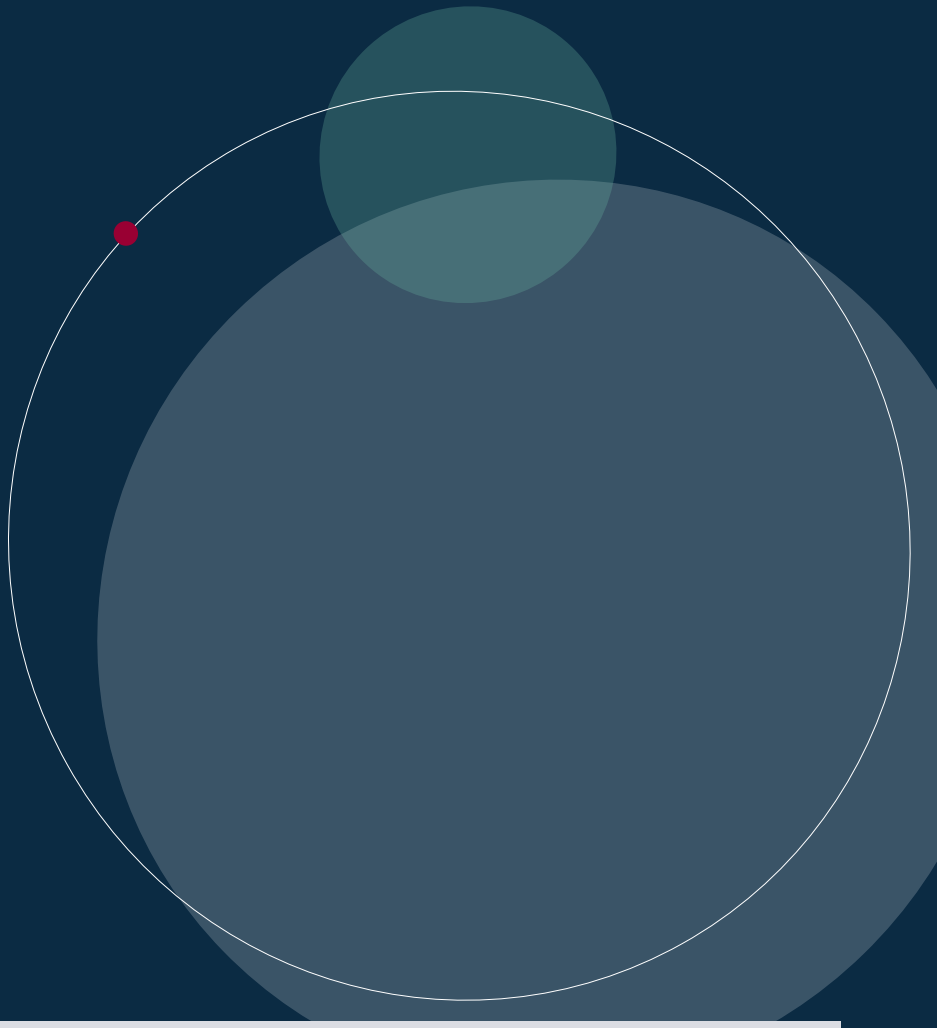
Except AI systems performing an assistive function for standard editing; that do not substantially alter the input data provided; authorised by law to detect / prevent, investigate / prosecute criminal offences

Emotion recognition systems / Biometric categorisation systems

Except AI systems permitted by law to detect / prevent / investigate criminal offences

AI systems generating or manipulating image, audio or video content constituting a deep fake / that generate or manipulate text published for information purpose on matters of public interest

Except AI systems authorised by law to detect / prevent / investigate / prosecute criminal offences, or where the AI-generated content has undergone a process of human review / editorial control under the editorial responsibility of an individual



Notes:

- The obligations provided for "AI systems subject to specific transparency obligations" may apply in addition to the obligations foreseen for High-risk AI systems.
- Further clarification from regulators will be necessary regarding the notion of "AI Systems intended to interact directly with natural persons" as it can be understood very broadly and can potentially capture a wide range of AI Systems.

GPAI Models

- **GPAI Model**: AI model that displays significant generality and is capable of competently performing 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.**

- A GPAI model will be classified as a **GPAI model with systemic risk** and be subject to **additional obligations** if it meets either of the following conditions:
  - **it has high impact capabilities** ➔ Presumption of high impact capabilities when the cumulative amount of computation used for its training measured in floating point operations is greater than 10^{25}; or
  - it **has capabilities / impact equivalent to a GPAI model with high impact capabilities** (based on the number of parameters; quality/size of the data set; amount of computation used for training the model; etc.).

A provider of a GPAI model that meets one of the conditions to be classified as GPAI model with systemic risk may argue that the GPAI model does not present, due its specific characteristics, systemic risks and should therefore not be classified as such.

Much more to explore!

Follow our EU AI Act Decoded Series as we delve into the intricacies of the EU AI Act.

https://www.linkedin.com/showcase/ai-data-digital/

Contact us

Anne-Gabrielle Haie
Partner in Steptoe's AI, Data & Digital practice

Notes:

- AI models that are used for research, development or prototyping activities before they are placed on the market are excluded from the definition of GPAI model, and thus they are not subject to the obligations applicable to GPAI models.
- Providers of GPAI model meeting one of the conditions to be classified as GPAI model with systemic risk must notify the European Commission without undue delay and in any event within 2 weeks after the condition is met / it becomes known that the condition will be met.
- The European Commission may decide, at its own volition or following a qualified alert form the scientific panel, to classify a GPAI model as a GPAI model with systemic risk.
- The list of GPAI models with systemic risk will be published.
- Thresholds, tools and benchmarks used to assess whether a GPAI model has high-impact capabilities or capabilities/impact equivalent to GPAI models with high impact capabilities can be amended by the European Commission in light of evolving technological developments.