White Label Consultancy

Unpacking AI Act

August 2024

Unpacking AI Act

Overview

The AI Act is structured to regulate AI comprehensively across all sectors. It categorises AI systems (AIS) based on the level of risk associated with their intended use, with the most stringent obligations placed on what is defined as high-risk AIS (HRAIS). Additionally, the AI Act regulates general-purpose AI (GPAI) models, with obligations tiered according to the systemic risk these models may pose.

The AI Act is designed to regulate AI providers established within the EU and in third countries. It affects deployers if they operate within the EU. To prevent circumvention, the AI Act also applies to deployers established in third countries if the output produced is intended for use in the EU. For example, this applies when an EU operator contracts services for activities classified as high-risk under the AI Act.

Similar to the GDPR and other recent EU digital regulations, the AI Act imposes substantial fines, which can reach up to EUR 35 million or 7% of the global annual turnover, depending on which amount is greater.

Timeline. When will the AI Act start to apply?

The AI Act will be implemented in phases:

1. February 2025. The first phase will introduce prohibitions on certain AI practices and mandate AI literacy requirements. Although regulatory bodies to oversee and enforce the AI Act are not expected until later in time (establishment planned by August 2, 2025), the AI Act anticipates prohibitions to positively impact other areas, such as civil law procedures.

2. August 2025. The second phase will introduce requirements for GPAI.

3. August 2026. The majority of the rules concerning AIS/HRAIS will come into effect.

Grace period:

The AI Act mandates that HRAIS already on the market or in service before its application will only be subject to the new rules if they undergo substantial modification after August 2026. Exceptionally, operators of HRAIS used by public authorities must comply with the AI Act by August 2030, regardless of substantial design or purpose modifications.

Prohibitions

The AI Act will ban certain AIS that are considered to pose an unacceptable risk. Specifically, the AI Act will prohibit AIS that manipulates or exploits real-time remote biometric identification in public spaces for law enforcement purposes, except in narrowly defined situations. It will also prohibit AIS that compile facial recognition databases through the untargeted scraping of images from the internet or CCTV footage and those that perform biometric categorisation to deduce sensitive data about individuals. There is a forbidden use case that can affect multiple organisations. In this regard, companies and educational institutions should be aware that the mere use of AIS to infer the emotions of employees or students is prohibited.

AI literacy requirements

AI literacy, set to be required from Q2 2025, aims to equip providers, deployers, and affected persons with the essential knowledge required to make informed decisions about AIS. This includes understanding the technical elements during the development phase, appropriate measures during use, methods for interpreting outputs, and the impact of AI-supported decisions on individuals. The European AI Board is anticipated to assist the Commission in advancing AI literacy tools and public awareness. Despite these measures, the primary responsibility ultimately resides with AI operators.

GPAI requirements

Providers of GPAI models, also known as "foundation models", will be required to prepare the following:

- technical documentation,
- policy to comply with copyright and related rights (Directive EU 2019/790),
- summary of the data used for training.

The minimal set of elements to be included in the technical documentation is determined in Annex XI. It includes a description of the GPAI model's intended use, applicable acceptable use policies, architecture, licence, design specifications, and training process details. The training data summary must be sufficiently detailed and publicly available. The logic behind the latter requirement is to facilitate copyright holders to exercise their rights; thus, the summary should be generally comprehensive in scope but not technically detailed to endanger trade secrets or confidential information. Providers are meant to list the main data collections or sets that went into training the model, such as large public or private databases or data archives, accompanied by a narrative explanation about other data sources used. The AI Office is expected to offer a template for this summary.

GPAI models will be considered models "with systemic risks" if the cumulative amount of computation power used for training is greater than 10^{25} floating point operations (FLOPs). To date, GPT-4 and Gemini Ultra surpass this threshold. Providers of such models are required to identify and mitigate these risks and ensure adequate cybersecurity protection. Additionally, for GPAI models deemed to have systemic risks, risk-management policies must include provisions for post-marketing monitoring and the reporting of serious incidents.

AIS requirements

The AI Act applies to AIS, which are broadly defined by their key characteristic: the ability to infer outputs, such as predictions, content, recommendations, or decisions, that can influence physical and virtual realities and to derive models or algorithms, or both, from inputs or data.

The AI Act follows a risk-based approach and categorises AIS into two groups: AIS and HRAIS. Typically, HRAIS are identified by reference to specific areas pre-defined in annexes to the AI Act, which the Commission plans to update annually.

However, derogations are possible. An AIS is not considered HRAIS if it does not significantly influence decision-making or harm protected legal interests. This is the case when the AIS performs narrow procedural tasks, enhances outcomes of previously completed human activities, detects patterns without replacing human assessment, or prepares data for further assessment. Examples include AIS that classify incoming documents, detect duplicates, improve language by aligning text to a certain brand messaging, flag potential anomalies, or provide solutions for translations. However, no derogations are possible in certain areas. For instance, all AIS in critical infrastructure are classified as HRAIS, as are all AIS used in remote biometric identification due to their reliance on sensitive data.

The requirement to train staff and increase AI literacy applies to all AIS operators, regardless of whether AIS is high-risk or not. AIS not deemed high risk may be voluntarily registered in the EU database to be created by the Commission. AIS classified as non-HRAIS due to a derogation still must register themselves in the EU database.

Furthermore, AIS that are not HRAIS but are intended to interact with people or generate content are subject to specific transparency obligations due to risks of impersonation or deception. Specifically:

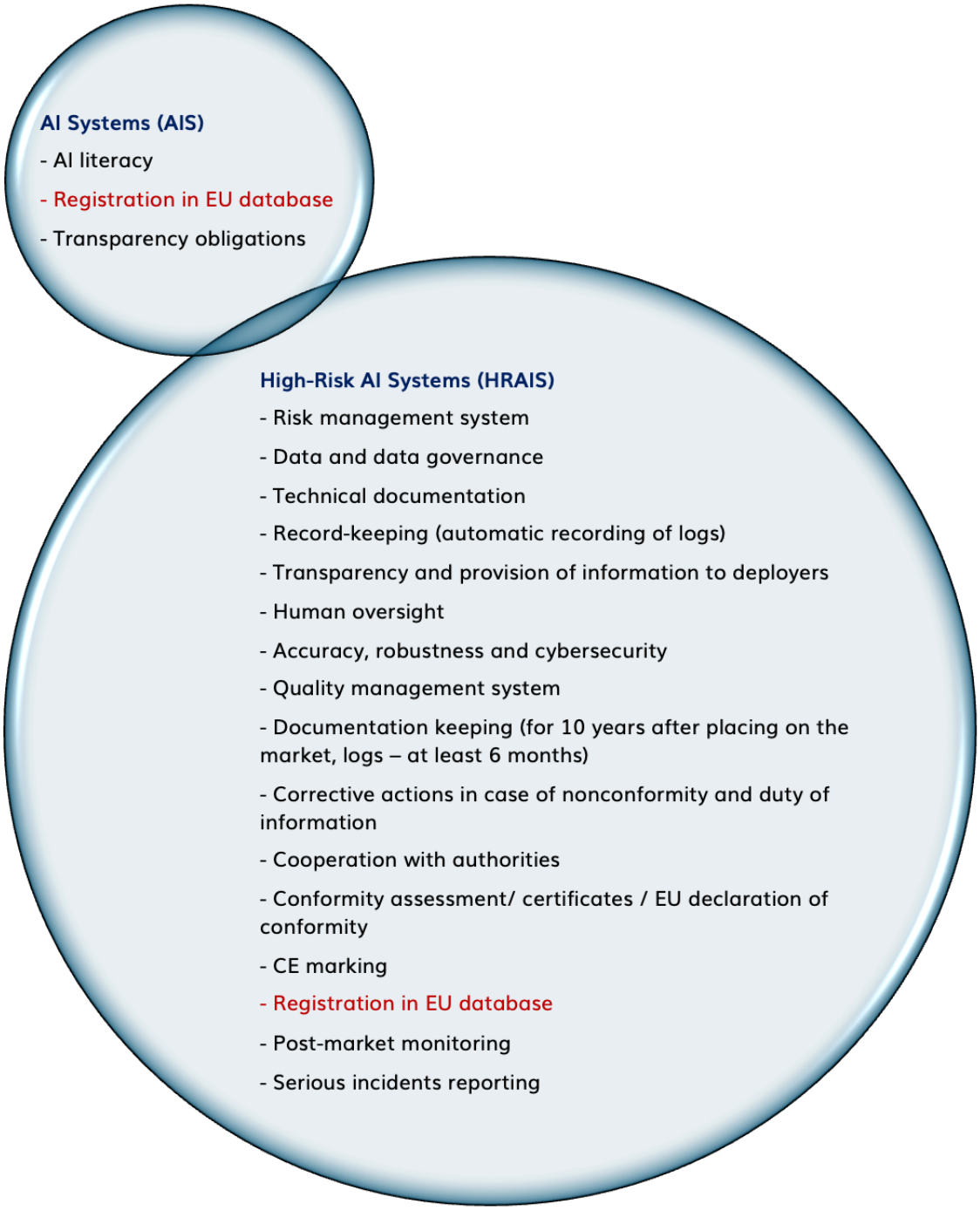
- Individuals must be notified that they are interacting with an AIS unless it is obvious from the perspective of a user "who is reasonably well-informed, observant and circumspect taking into account the circumstances and the context of use".
- Individuals must also be informed when exposed to AIS, which processes biometric data to identify or infer emotions or intentions or that categorise them based on attributes like eye colour, tattoos, personal traits, preferences, or interests.
- AIS-generating synthetic content must include in a machine-readable format indicating that the output was generated or manipulated by AI, not a human.

Finally, the AI Act clarifies that even non-high-risk AIS must maintain safety safeguards, with Regulation EU 2023/988 acting as a supplementary measure to ensure the general safety of products made available on the EU market.

HRAIS requirements

HRAIS includes systems used in areas listed in Annex III of the AI Act, such as systems to be used for remote biometric identification or to safely operate critical infrastructure. Additionally, any AIS that serves as a safety component of a regulated product or is itself a regulated product will also qualify as high risk. Annex I of the AI Act outlines the relevant EU legislation covering products like medical devices and cars.

Compared to general AIS, the list of obligations placed for HRAIS is considerably more extensive.



HRAIS

Annex I

- **machinery** *(Directive 2006/42/EC)*
- **toys** *(Directive 2009/48/EC)*
- **recreational craft and personal watercraft** *(Directive 2013/53/EC)*
- **lifts and safety components for lifts** *(Directive 2014/33/EC)*
- **equipment and systems intended for use in potentially explosive atmospheres** *(Directive 2014/34/EC)*
- **radio equipment** *(Directive 2014/53/EC)*
- **pressure equipment** *(Directive 2014/68/EC)*
- **cableway installations** *(Regulation EU 2016/424)*
- **personal protective equipment** *(Regulation EU 2016/425)*
- **appliances burning gaseous fuels** *(Regulation EU 2016/426)*
- **medical devices** *(Regulation EU 2017/745)*
- **in vitro diagnostic medical devices** *(Regulation EU 2017/746)*
- **civil aviation, including unmanned aircraft and their engines, propellers, parts, and equipment to control them remotely** *(Regulation EC 300/2008; Regulation EU 2018/1139)*
- **two and three-wheel vehicles and quadricycles** *(Regulation EU 168/2013)*
- **agricultural and forestry vehicles** *(Regulation EU 167/2013)*
- **marine equipment** *(Directive 2014/90/EU)*
- **rail system** *(Directive EU 2016/797)*
- **motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles** *(Regulation EU 2018/858; Regulation EU 2019/2144)*

Products or their safety components that fall under EU legislation listed in Annex I have the flexibility to choose how to ensure compliance. This means that the required processes can be integrated into existing documentation.

Regarding conformity assessments, Annex I HRAIS operators can undertake these through regulatory bodies that are specific to their sectors, such as medical devices, in vitro diagnostic devices, cars, toys, etc. The AI Act stipulates that only one conformity assessment is required, and this should be the one already mandated by sector-specific laws. The AI Act is designed to not alter the specific logic, methodology, or overall structure of conformity assessments as prescribed by these sectoral laws.

Consistent with the established concept of substantial modification, an AIS must undergo a conformity reassessment when there are changes that could impact compliance (such as changes to the operating system or software architecture) or when there is a change in the intended purpose of the system.

HRAIS

Annex III

- **biometrics** (remote biometric identification, biometric categorisation, emotion recognition)
- **critical infrastructure** (digital infrastructure, road traffic, supply of water, gas, heating or electricity)
- **education and vocational training** (systems determining access or admission, evaluating learning outcomes, assessing education level access, monitoring prohibited behaviour during tests)
- **employment, workers' management and access to self-employment** (systems used for recruitment, placing targeted advertisements, analysing and filtering applicants, evaluating candidates, affecting terms of work-related relationships, promotion, termination, task allocation, performance evaluation)
- **access to essential private and public services and benefits** (public assistance, including healthcare, creditworthiness (exception: fraud detection), assessment in the case of life and health insurance, emergency calls evaluation and classification)
- **law enforcement**
- **migration, asylum and border control management**
- **administration of justice and democratic processes**

The initial review of the AI Act is scheduled for August 2, 2029, with subsequent reviews every four years after. However, the Commission intends to annually review the list of HRAIS.

Providers of HRAIS are subject to extensive requirements, as outlined before. These include establishing a risk management system that encompasses risk identification, testing, and the implementation of risk mitigation measures. The Commission has identified the quality management system requirement as the most burdensome and costly; thus, microenterprises will be allowed to meet this requirement in a simplified way. The Commission is also in the process of developing guidance to clarify which elements can be simplified. Third-party conformity assessments are expected to be developed, and notified bodies will be established by August 2, 2025. Initially, operators of Annex III HRAIS, except those processing biometric data, will only need to conduct a conformity self-assessment based on internal controls and will not be required to involve a notified body.

Additionally, HRAIS must display the CE marking (whether physical or digital, as appropriate) and must register in an EU database that the Commission will establish. The AI Act imposes on HRAIS providers responsibilities of post-market monitoring and reporting serious incidents, including malfunctions that lead to death or serious health damage, severe disruptions to critical infrastructure, violations of fundamental rights, or significant damage to property or the environment.

Enforcement of AIS/HRAIS rules will primarily be handled at the national level. Each EU member state must designate the appropriate authority within one year of the AI Act enactment.

Authors

Nicholai Pfeiffer

Managing Partner
White Label Consultancy
np@whitelabelconsultancy.com

Arina Kostina

Consultant, Data Protection
White Label Consultancy
ako@whitelabelconsultancy.com

Federico Marengo

Senior Consultant
White Label Consultancy
fma@whitelabelconsultancy.com

Przemysław Gruchała

Senior Consultant
White Label Consultancy
pgr@whitelabelconsultancy.com

Contact

hello@whitelabelconsultancy.com
+45 71 74 74 54

NORWAY

Main Office Oslo

+47 4141 2168

Fjordalléen 16
0250 Oslo

DENMARK

Copenhagen Office

+45 71 74 74 65

Dampfærgevej 27
2100 København Ø

POLAND

Warsaw Office

+48 515 07 99 77

Ul. Marszałkowska 58/15
00-545 Warszawa

UNITED ARAB EMIRATES

Dubai Office

+971 50 4536616

Dubai World Trade Center
Dubai

WHITELABELCONSULTANCY.COM