Cracking the Code - Exploring AI Act for High-Risk AI Systems

As days go by, artificial intelligence ("**AI**") continues to delve into our daily life. It is one of the reasons determining the representatives of the European Union ("**EU**") to become pro-active in regulating its use, in particular in relation to AI systems. In this context, on July 12th, 2024, the long-awaited *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)*[1] (the "**AI Act**" or the "**Regulation**") was published in the Official Journal of EU.

The AI Act introduces a new risk-based approach that will uniformly apply across all Member States and establishes a comprehensive framework in relation to AI systems, including high-risk AI systems, prohibited AI practices and general-purpose AI models.

Regarding the next steps, the AI Act will enter into force on August 1st, 2024, and it will be fully applicable from August 2nd, 2026 (with certain exceptions).

1. What is an AI system? How is it different from a high-risk AI system?

The AI Act proves to be an intricate piece of legislation in what regards high-risk AI systems (further referred to as "**HRS**"), and we intend to shed some light in this respect, aiming to help entities to further plan their respective conformity strategies.

For providers[2] (who are the developers or suppliers of HRS) and deployers[3] (users of HRS), as well as importers[4] and distributors[5], the stakes of qualifying a system as HRS are rather significant, since, by way of example, providers are required to adhere to **rigorous standards to guarantee the trustworthiness, transparency, and accountability of their systems.**

Perhaps the most important step from which any operator[6] should proceed in contemplating its conformity actions is the definition of the AI system itself, which was intentionally left as broad ranging to ensure that the AI Act does not become easily outdated:

*"**AI system** means a machine-based system that is 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."*

Furthermore, in order to qualify as HRS, the AI system must fulfil several requirements or must be part of a list of pre-defined areas expressly referred to, as per the following three situations:

(i) TYPE 1 HRS: the AI systemis intended to be used as a **safety component of a product** (*i.e.,* a component of a product or of an AI system which fulfils a safety function for that product or AI system, or the failure or malfunctioning of which endangers the health and safety of persons or property);

(ii) TYPE 2 HRS: the AI system **is itself a product**, covered by the Union harmonization legislation listed in **Annex I** (in particular, such products are machinery, toys, lifts, equipment and protective systems intended for use in potentially explosive atmospheres, radio equipment, pressure equipment, recreational craft equipment, cableway installations, appliances burning gaseous fuels, medical devices, and *in vitro* diagnostic medical devices, automotive and aviation);

NB! TYPE 1 HRS and TYPE 2 HRS are required to undergo the **conformity assessment procedure with a third-party conformity assessment body**, prior to their placing on the market or putting into service.

(iii) TYPE 3 HRS: the AI system is the one referred to in **Annex III** (such as biometrics[7], AI systems intended to be used as safety components in critical infrastructure, education and vocational training, employment, evaluating the creditworthiness of natural persons or establish their credit score, law enforcement, migration and administration of justice and democratic processes)**.**

NB! A distinction must be made between biometrics as a prohibited practice and biometrics as HRS. More specifically, the following represents cases where biometrics are considered prohibited practice:

•€€€€€€€€ categorizing individuals based on sensitive biometric data without a lawful purpose, in order to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in law enforcement;

•€€€€€€€€ employing "real-time" biometric identification systems in public spaces for law enforcement, except in specific circumstances such as preventing serious crimes or locating missing persons. In principle, relying on such an exception will require thorough assessments, technical and organizational measures, notifications, and a warrant;

•€€€€€€€€ untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases;

•€€€€€€€€ using AI systems intended to detect the emotional state of individuals in situations related to the workplace and education[8] except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons.

Nevertheless, if the biometrics in question do not fall into the categories above, they could be classified as HRS and be subject to the relevant regulation. This is the case, for example, for:

•€€€€€€€€ remote biometric identification, meaning an AI system used for the purpose of identifying natural persons, without their active involvement, typically at a distance (excluding certain AI systems intended to be used for biometric verification, including authentication, whose sole purpose is to confirm that a specific natural person is who that person claims to be);

•€€€€€€€€ AI systems intended to be used for biometric categorization (*i.e.* such specific categories can relate to aspects such as sex, age, hair and eye colour, tattoos, behavioural or personality traits, language, religion *etc.*);

•€€€€€€€€ AI systems intended to be used for emotion recognition, other than in the area of the workplace and education institutions (these systems refer to emotions or intentions such as happiness, sadness, anger, surprise, embarrassment, excitement, shame or amusement, but they do not include physical states, such as pain or fatigue).

AI systems intended to be used solely for the purpose of enabling cybersecurity and personal data protection measures should not qualify as HRS.

Finally, **profiling[9] of natural persons always qualifies an AI system as HRS**.

Guidelines detailing the practical implementation of AI system classification, along with a comprehensive list of practical examples of use cases of AI systems that are high-risk and not high-risk, are anticipated to be released no later than 18 months after the AI Act comes into force.

2. Are there any derogations from HRS qualification?

Even if included in the list of TYPE 3 HRS, operators may prove that the AI system poses no significant risk to health, safety, or fundamental rights, and does not materially influence decision-making outcomes. These exemptions include:

•€€€€€€€€ *performing narrow procedural tasks* (this AI system can be used in a variety of areas, such as HR or customer management, to detect duplicates in many applications or to classify documents into categories);

•€€€€€€€€ *improving previous human activities* (*e.g.* AI systems used to improve the language used in documents already drafted, in order to be more business-like or to be written in a certain style of academic language);

•€€€€€€€€ *detecting decision-making patterns without replacing human assessment* (*e.g.* an AI system that, given a certain grading pattern of a teacher, can be used to check ex post whether the teacher may have deviated from the grading pattern so as to flag potential inconsistencies or anomalies);

•€€€€€€€€ *performing preparatory tasks* (such as smart solutions for file handling, including functions like indexing, searching, text and speech processing, or linking data to other sources, or AI systems used for translation of initial documents).

To ensure traceability and transparency, a provider who considers that an AI system is not high-risk based on those conditions should prepare the documentation of the assessment before that system is placed on the market or put into service and should provide this documentation to national competent authorities upon request. However, such a provider is still required to **register the AI system in the EU database for HRS**.

3. What are the requirements related to HRS?

To mitigate the risks resulting from HRS placed on the market or put into service and to ensure a high level of trustworthiness, certain mandatory requirements[10] apply in relation to HRS.

The AI Act has established a sizeable number of requirements for different types of roles, starting from inception up to the whole HRS supply chain. Some of the most important obligations are outlined below, divided according to the subjects who must abide them.

3.1. What are the HRS provider's specific obligations?

HRS providers are subject to specific requirements, including extended responsibility for products containing an AI system, notwithstanding the requirements deriving from other pieces of Union harmonized legislation.

According to the AI Act, among other things, HRS providers must:

a) establish, implement, document, and maintain a risk management system to identify and mitigate potential risks;

b) ensure data governance by using high-quality data sets;

c) create and update technical documentation before the HRS is placed on the market or put into service;

d) design the HRS for automatic event recording throughout their lifecycle, facilitating the identification of potential risks and substantial modifications;

e) prioritize transparency in the design and development of the HRS, ensuring that deployers can understand and use system outputs appropriately;

f) design and develop the HRS in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which they are in use;

g) design the HRS to achieve appropriate levels of accuracy, robustness, and cybersecurity and to perform consistently in those respects throughout their lifecycle;

h) establish a comprehensive quality management system;

i) keep specific documentation available to competent authorities for a period ending 10 years after the HRS has been placed on the market or put into service;

j) take corrective actions if non-compliance is suspected (including bringing the AI system into conformity, withdrawing it, disabling it, or recalling it);

k) cooperate with the competent authorities, as requested;

l) comply with registration obligations in the EU database for HRS;

m) appoint an authorized representative (if established in third countries).

3.2. What are the HRS deployer's specific responsibilities?

While deployers of HRS have fewer obligations compared to providers, they still play a crucial role in ensuring compliance and mitigating risks. More specifically, deployers must:

a) use the HRS in accordance with the instructions provided by the providers, ensuring proper data usage and system monitoring;

b) if exercising control over the input data, ensure that such data is relevant and sufficiently representative for the intended purpose of the HRS;

c) assign human oversight to natural persons who have the necessary competence, training, and authority, as well as the necessary support;

d) inform providers and relevant authorities promptly if they suspect risks or identify incidents related to system use and suspend system usage;

e) retain automatically generated system logs, for a period appropriate to the intended purpose of the HRS, of at least six months, unless provided otherwise by the Union or national law;

f) before putting into service or using an HRS at the workplace, deployers who are employers shall inform workers' representatives and the affected workers that they will be subject to the use of the HRS.

g) inform natural persons that they are subject to the use of the HRS;

h) cooperate with competent authorities in implementing regulatory requirements;

i) when applicable, conduct data protection impact assessments and obtain authorization for post-remote biometric identification system use, adhering to strict limitations on data processing and documentation requirements;

j) when applicable, complete a fundamental rights impact assessment before deploying the AI system (for certain deployers, including public bodies and private entities providing public services, such as banks, insurers, hospitals, and schools).

3.3. What are the HRS importers' specific obligations?

Before placing the HRS on the market, importers must verify that the AI system has undergone the necessary conformity assessment, that technical documentation meets regulatory standards, and that it carries the CE marking[11], the EU declaration of conformity, and instructions for use. If there are suspicions of non-compliance or falsification, importers are obligated to refrain from placing the system on the market and inform relevant parties and authorities promptly.

Additionally, importers shall indicate their name, registered trade name or registered trademark, and the address at which they can be contacted on the HRS and on its packaging or its accompanying documentation, where

applicable.

3.4. What are the HRS distributors' specific obligations?

Among other responsibilities, distributors of HRS are tasked with verifying that the AI system bears the CE marking, includes the EU declaration of conformity and instructions for use, and that providers and importers have complied with their respective obligations. If non-compliance is suspected, distributors will withhold market availability and promptly inform providers or importers of any risks.

Furthermore, distributors must maintain proper storage conditions and take corrective actions if non-compliance is detected, cooperating with authorities, and providing necessary information upon request to demonstrate conformity of the AI system with the requirements set out in the AI Act.

3.5. Are there any particular obligations related to AI systems that pose the risk of impersonation or deception?

Irrespective of whether they qualify as HRS[12] or not, the AI Act regulates a dedicated chapter (*i.e.* Chapter IV) on transparency obligations for providers and deployers of certain AI systems that may pose specific risks of impersonation or deception (*e.g.* ChatGPT-based systems or systems that generate *deep fake* content).

Among these requirements, **providers** must ensure that individuals interacting *directly* with AI systems are informed about their nature, except where this is obvious or for certain authorized criminal justice functions. Moreover, if generating synthetic content (*e.g.* an image that has been fully or partially created using computer-generated graphics), they must ensure that the outputs are clearly marked as *artificially generated or manipulated*, with exceptions for standard editing purposes or authorized criminal justice uses.

Deployers of emotion recognition or biometric categorization systems must inform individuals about system operations and adhere to data regulations, except for AI systems used for biometric categorization and emotion recognition that are permitted by law to detect, prevent, or investigate criminal offences.

For AI systems creating *deep fake*[13] content, deployers must disclose its artificial nature, except for authorized criminal justice purposes[14]. Similarly, deployers of AI-generated text for public information must disclose that the text has been *artificially generated or manipulated* (with certain exceptions). These disclosures must be clear and provided during the initial interaction, meeting accessibility requirements.

4. Could the responsibility of any of the above roles be extended?

It should be noted that, under certain specific conditions, any distributor, importer, deployer or other third-party may be considered as a provider of an HRS and therefore assume all the relevant obligations, when:

(i) they affix their name or trademark to an existing HRS,

(ii) substantially modify such a system, or

(iii) change the intended purpose of an AI system, making it a high-risk one.

In such cases, the initial provider waives responsibility for the specific AI system, but must cooperate[15] with the new providers, supplying them with the necessary information and technical access, as well as providing other necessary assistance for their full compliance.

Furthermore, for TYPE 1 HRS included in Section A of Annex I (relating, *inter alia*, to machinery, safety of toys, lifts and safety components for lifts, radio equipment, cableway installations, medical devices and in vitro diagnostic medical devices), the product manufacturer shall assume provider obligations if the AI system is marketed or put into service under their name or trademark.

5. Conclusions

The AI Act establishes a wide-ranging framework for regulating HRS, aiming to balance innovation with the protection of fundamental rights and human safety. By imposing obligations on both providers and deployers, and by emphasizing transparency and oversight, this piece of legislation aims to foster trust in AI technologies while mitigating potential risks.

Given the increasingly ample use of AI technologies in most of the companies today, as well as the complexity of the AI Act altogether, putting in place strategies aimed at ensuring conformity becomes a necessity as we speak. In fact, the AI Act itself states that HRS providers are encouraged to start complying with the relevant obligations, on a voluntary basis, as early as during the transitional period.

---

1. Available here.

*2. A **provider** 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 it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge.*

*3. A **deployer** is a 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.*

*4. An **importer** is a natural or legal person located or established in the Union that places on the market an AI system that bears the name or trademark of a natural or legal person established in a third country.*

*5. A **distributor** is a natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available*

on the Union market.

*6. An "**operator**" means a provider, product manufacturer, deployer, authorized representative, importer or distributor.*

7. Insofar as their use is permitted under relevant Union or national law.

8. Considering the imbalance of power in the context of work or education, combined with the intrusive nature of these systems, such systems could lead to detrimental or unfavourable treatment of certain natural persons or whole groups thereof. Therefore, the placing on the market, the putting into service, or the use of AI systems intended to be used to detect the emotional state of individuals in situations related to the workplace and education are prohibited.

9. Profiling means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.

10. It is important to note that most of these requirements will be effective after the AI Act becomes fully applicable. However, obligations on HRS TYPE 1 and HRS TYPE 2 shall apply from August 2nd, 2027.

11. Which is a marking used by the provider to indicate that an AI system is in conformity with the requirements set out in the AI Act.

12. If also qualified as HRS, the use of such systems should therefore be subject to both specific transparency obligations, without prejudice to the requirements and obligations for HRS.

13. The AI Act defines what should be considered deep fake, meaning an AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful.

14. However, for artistic, creative, or fictional works, they only need to disclose the existence of manipulated content in a non-disruptive manner.

15. This obligation shall not apply in cases where the initial provider has clearly specified that its AI system is not to be changed into an HRS and therefore it does not fall under the obligation to hand over the documentation.