Zooming in on AI – EU AI Act – What are the obligations for limited-risk AI systems?

11 November 2024

The EU Artificial Intelligence Act ("AI Act") exemplifies a highly advanced risk-based approach to European regulation. One of its distinguishing features is the detailed classification of various risk levels associated with AI technology.

Each level comes with its own set of requirements. As detailed in article #10 of our 'Zooming in on AI' series: What are the obligations for "high risk AI systems"?, high-risk AI systems undergo stricter scrutiny to mitigate potential dangers.

This blog post delves into the requirements for another risk category – i.e., limited-risk AI systems, as defined by the AI Act, which benefit from a lighter regulatory touch.

Limited-risk AI systems

Certain AI systems intended to interact with individuals or to generate content may pose specific risks of impersonation or deception. Such AI systems are classified as limited-risk AI systems under the EU AI Act (i.e., AI systems with transparency risks). These systems are unlikely to cause significant harm or violate fundamental rights

AI systems with transparency risks include those that:

1. interact directly with individuals (such as chatbots and digital assistants);
2. generate synthetic audio, image, video, or text content (such as ChatGPT);
3. generate or manipulate image, audio or video content resulting in a deep fake;
4. generate or manipulate text intended to inform the public on matters of public interest; and
5. are an emotion recognition system (e.g. those used in the entertainment industry) or a biometric categorization system (e.g. facial recognition)[1] .According to EU lawmakers, in certain circumstances, the use of AI systems with transparency risks should be subject to specific transparency obligations.

Transparency obligations for limited-risk AI systems are different for providers and deployers of such systems. We discuss the interplay between "deployers" and "providers" in our Zooming in on AI – #4: What is the interplay between "Deployers" and "Providers" in the EU AI Act? blog post.

Provider obligations

According to Articles 50(1) and 50(2) of the AI Act, providers of AI systems with transparency risks must adhere to the below requirements:

1. **Chatbots and digital assistants:** Providers must ensure that AI systems intended for direct interaction with individuals are designed and developed to inform those individuals that they are engaging with an AI system. This requirement is waived only if it is evident to a reasonably well-informed, observant, and cautious person, considering the circumstances and context of use. This obligation does not apply to AI systems legally authorized to detect, prevent, investigate, or prosecute criminal offenses, provided that appropriate safeguards for the rights and freedoms of third parties are in place, unless these systems are accessible to the public for reporting criminal offenses.
2. **AI-generated synthetic content:** Providers of AI systems, including those that generate synthetic audio, image, video, or text content, must ensure that the outputs of these systems are marked in a machine-readable format and are detectable as artificially generated or manipulated. These providers must ensure that their technical solutions are:

- **Effective**: The solutions should reliably mark and detect synthetic content.
- **Interoperable**: The solutions should work seamlessly across different platforms and systems.
- **Robust and reliable**: The solutions should be dependable and maintain their effectiveness over time.

These requirements should be met as far as technically feasible, considering: (i) the specific characteristics and limitations of various types of content; (ii) the costs associated with implementation; and (iii) the generally acknowledged state of the art, as reflected in relevant technical standards.

The above transparency obligations will not apply to AI systems that perform an assistive function for standard editing or that do not significantly alter the input data provided by deployers, nor the semantics of that data.

Deployer obligations

Articles 50(3) and 50 (4) of the AI Act specify transparency obligations for deployers of limited-risk AI systems:

1. **Deep fakes:** A deep fake is defined in the AI Act as AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful[2]. Deployers of AI systems that generate or manipulate image, audio, or video content resulting in deep fakes must disclose that the content has been artificially generated or manipulated. This obligation does not apply when the use is authorized by law for the purposes of detecting, preventing, investigating, or prosecuting criminal offenses. When the content is part of an evidently

artistic, creative, satirical, fictional, or similar work or program, the transparency obligations are limited. In such cases, disclosure of the existence of the generated or manipulated content must be made in an appropriate manner that does not interfere with the display or enjoyment of the work.

2. **AI generated or manipulated text:** Entities deploying AI systems to generate or manipulate text intended for public dissemination on matters of public interest must disclose that the content has been artificially generated or manipulated. This requirement does not apply in the following circumstances: (i) when the use of AI is authorized by law for the detection, prevention, investigation, or prosecution of criminal offenses; or (ii) when the AI-generated content has been subjected to human review or editorial control, and a natural or legal person holds editorial responsibility for the publication of the content.
3. **Emotion recognition system or biometric categorisation system:** Deployers of such systems must inform the individuals exposed to these systems about their operation. Additionally, they must process personal data in compliance with the EU General Data Protection Regulation (GDPR)[3], where applicable. However, this obligation does not extend to AI systems used for biometric categorization and emotion recognition that are legally authorized to detect, prevent, or investigate criminal offences. Such systems must still adhere to appropriate safeguards to protect the rights and freedoms of third parties and operate in accordance with Union law.

Timing and format of delivering required notices

Information regarding limited-risk AI systems must be provided clearly and distinguishably at the first interaction or exposure of the user with limited-risk AI system. When fulfilling this obligation, it is essential to consider the characteristics of individuals who belong to vulnerable groups due to their age or disability, especially if the AI system is designed to interact with these vulnerable groups. It is crucial that this information and these notifications are provided in formats that are accessible to persons with disabilities[4].

European Commission's role

The European Commission will review and potentially amend the list of limited-risk AI systems every four years[5]. The European Commission will also facilitate the creation of codes of practice (guidelines), to implement detection and labeling obligations for artificially generated or manipulated content. When issuing such guidelines, the Commission will pay particular attention to the needs of small and medium enterprises including start-ups, of local public authorities and of the sectors most likely to be affected by the AI Act[6].

Interplay with the transparency obligations under the GDPR and Digital Services Act (“DSA”)

When personal data is processed, the GDPR transparency requirements set forth in Articles 12-14 apply alongside the AI Act obligations, particularly regarding the purpose of data collection[7].

The transparency obligations imposed on providers and deployers of certain AI systems by the AI Act are also crucial for the effective implementation of the DSA[8]. These obligations are particularly pertinent for providers of very large online platforms or very large online search engines, as they must identify and mitigate systemic risks associated with the dissemination of artificially generated or manipulated content. The requirement to label AI-generated content under the AI Act does not also affect the obligation outlined in Article 16(6) of the DSA. Thus, the providers of hosting services must still process notices regarding illegal content received.

Enforcement and fines

National competent authorities will ensure compliance with the transparency requirements. Non-compliance can result in administrative fines of up to EUR 15 million or 3% of the operator's total worldwide annual turnover for the preceding financial year, whichever is higher.

Timeline

As detailed in our first article Zooming in on AI: When will the AI Act apply?, the transparency requirements for limited-risk AI systems under Article 50 of the AI Act will apply from 2 August 2026.

Footnotes

[1] See Article 50 of the AI Act.

[2] See Article 3(60) of the AI Act

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC

[4] See Article 50(5) and Recital 132 of the AI Act

[5] See Article 112(2)(b) of the AI Act

[6] See Article 96(1) of the AI Act

[7] See Art. 50(6) of the AI Act

[8] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)