

## TECHNOLOGY

# Implementation of transparency obligations for AI systems

*ALG's guide to Article 50 of the EU AI Act*

24 MIN READ

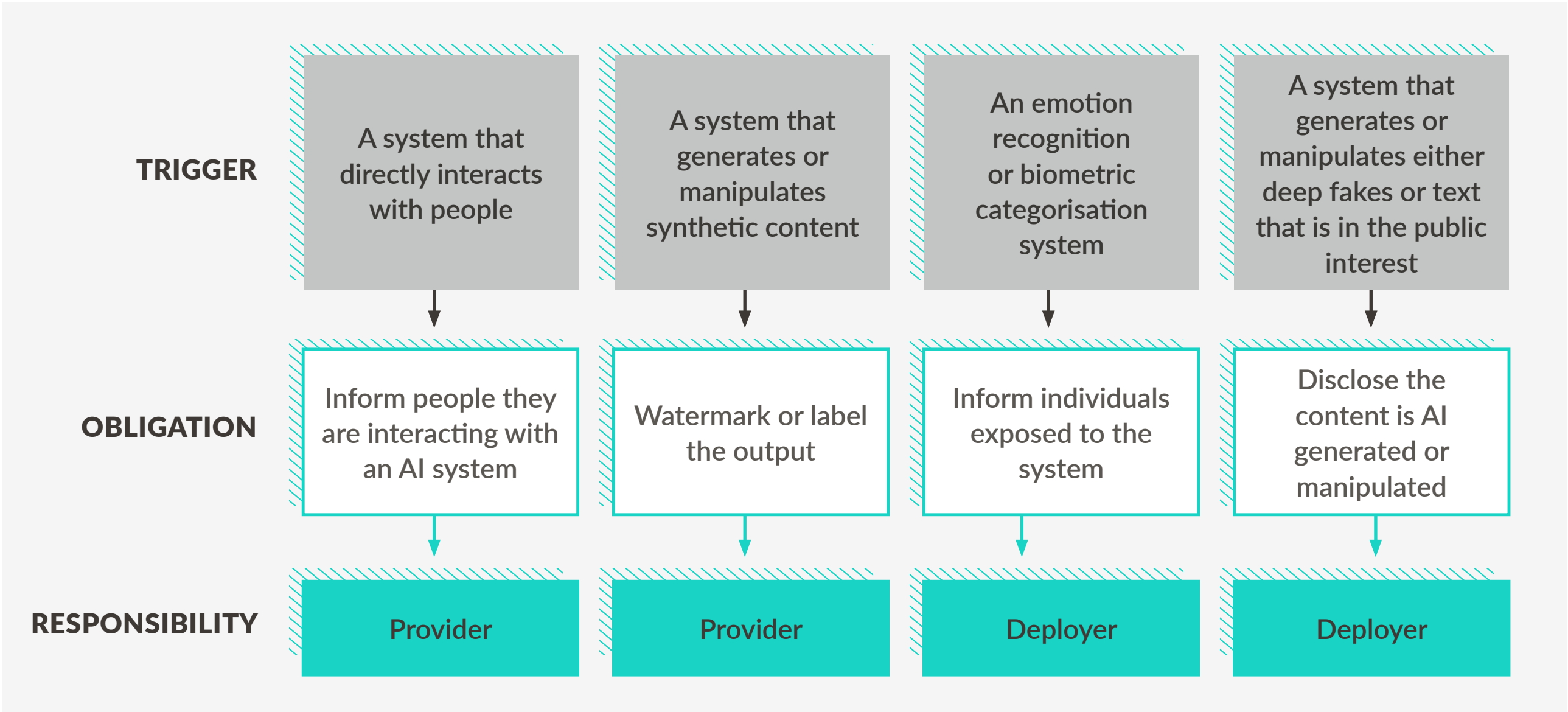
The EU AI Act establishes a comprehensive legal framework for the regulation of artificial intelligence. Article 50 of the AI Act sets out specific transparency obligations for providers and deployers of AI systems, including requirements to label AI-generated content, disclose interactions with AI, inform individuals exposed to emotion recognition systems, and mark deepfakes. Since 2 August 2026, organisations subject to the AI Act have been required to comply with these obligations.

In this guide, we examine the Article 50 transparency requirements in detail, consider the European Commission's recently published guidelines and voluntary Code of Practice on their interpretation, and set out the practical considerations for organisations looking to achieve and maintain compliance.



Key takeaways

- Article 50 of the AI Act sets out transparency obligations in respect of in-scope AI operators for specified uses of AI systems.
- These requirements are among the first substantive positive obligations under the AI Act to come into force. Since 2 August 2026, organisations are required to consider whether they are in-scope and what they need to do to comply.
- The obligations under Article 50 are divided into four distinct categories, each triggered by a specific condition:



- Two documents have been published by the European Commission to assist operators of AI systems understand and comply with their obligations:
  - » guidelines that break down the obligation triggers and the basics of the obligations, and
  - » a voluntary code of practice that helps signatory organisations demonstrate compliance with certain obligations by adhering to the code.





## Introduction

The EU Artificial Intelligence Act (the “**AI Act**” or the “**Act**”) establishes a comprehensive legal framework for the regulation of artificial intelligence across the EU. One of the first set of substantive positive obligations under the AI Act relating to AI transparency came into force on 2 August 2026.

In this guide, we take a look at the AI Act’s transparency obligations, when they apply, and the practical steps organisations can take to comply with them.

### What is Article 50 of the AI Act?

Article 50 of the AI Act imposes transparency obligations in relation to AI systems when they are used in one of the following four ways:

- the AI system is intended to interact directly with individuals
- the AI system generates synthetic content such as audio, image, video or text
- the AI system is an emotion recognition or biometric categorisation system, and/or
- the AI system generates or manipulates audio, image, or video content to create a deep fake or text that is published for the purpose of informing the public.

Being subject to, and complying with, the Article 50 transparency risk obligations under the AI Act does not negate requirements under the rest of the AI Act. Accordingly, it is possible for an AI system to be subject to both a transparency obligation and high-risk obligations under the AI Act.

### What is an AI system?

In order to be in scope of Article 50 of the AI Act, the system must fall within the definition of AI system under the AI Act. This means it must be:

- a machine-based system
- that is designed to operate with varying levels of autonomy, and
- infers from the input it receives how to generate outputs such as predictions, content recommendations or decisions.

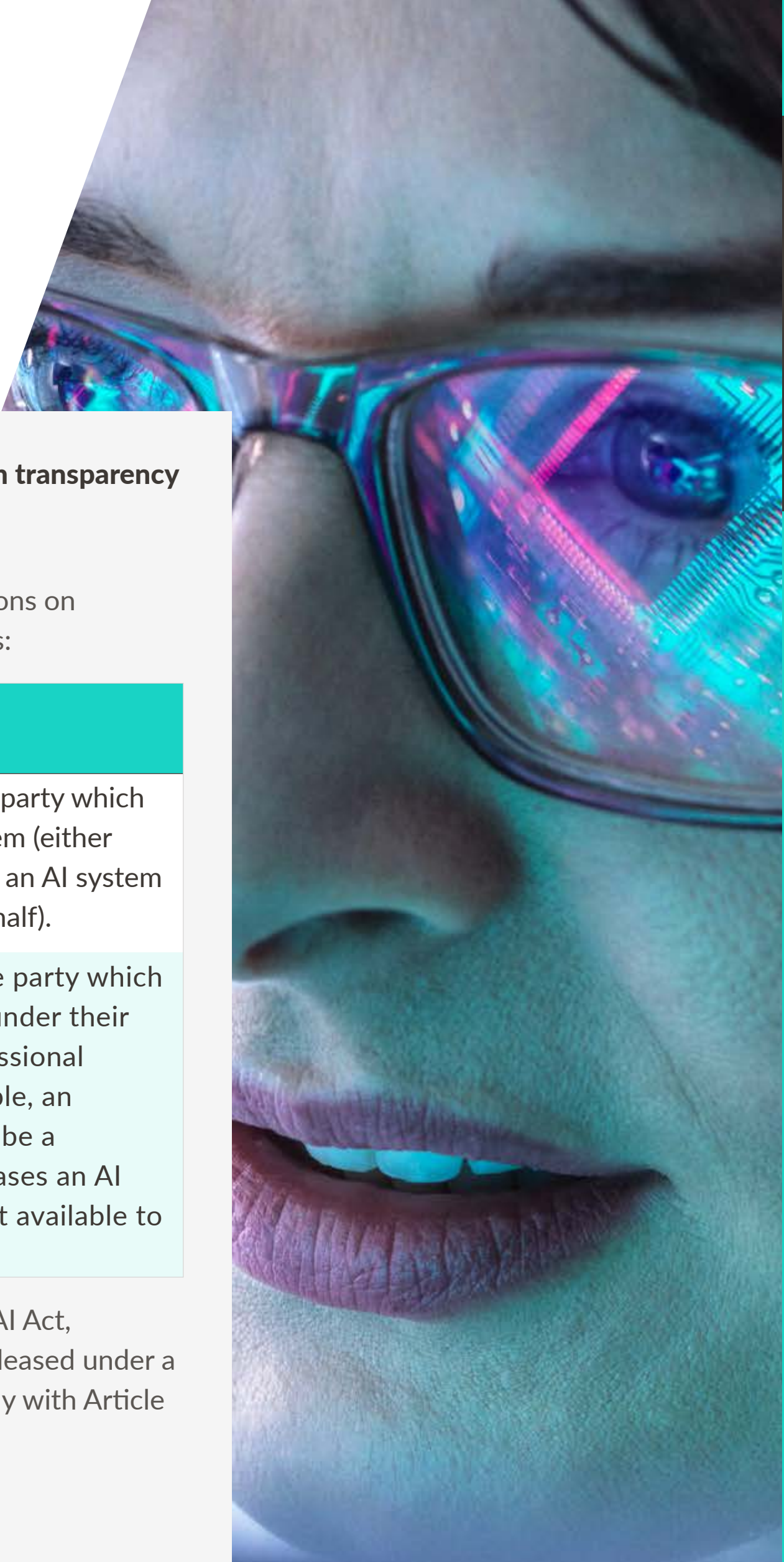
This definition is intentionally broad and expressly intended to cover a wide range of AI systems, including both complex generative AI tools and simpler systems utilising techniques such as text matching, knowledge-based responding, and decision trees.

### Article 50 scope: Who must comply with transparency obligations

Article 50 imposes transparency obligations on “providers” and “deployers” of AI systems:

| CATEGORY OF OPERATOR | DESCRIPTION                                                                                                                                                                                                                         |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Provider             | This is generally the party which <b>develops</b> an AI system (either directly or by having an AI system developed on its behalf).                                                                                                 |
| Deployer             | This is generally the party which <b>uses</b> an AI system under their authority in a professional capacity. For example, an organisation would be a deployer if it purchases an AI system and makes it available to its employees. |

Unlike with certain other aspects of the AI Act, providers and deployers of AI systems released under a free and open source licence must comply with Article 50 of the AI Act.





Exclusions from scope

- **Purely personal non-professional activity.** Article 50 does not apply to deployers who are natural persons using the AI system for “*purely personal, non-professional activity*”. Evidence of a regular economic benefit gained in connection with this activity or where this activity arises in a professional context (or on behalf of a deployer acting in a professional context) will tend against a finding that this exclusion applies. The Guidelines also emphasise that the exclusion attaches to the activity and not the system. The requirement is that a person should act in both a personal and a non-professional capacity to fall within the scope of the exception.
- **AI systems developed and used solely for scientific research and development.** Article 50 does not apply to AI systems (or their outputs) where the AI system is specifically developed and put into service for the sole purpose of scientific research and development. According to the Guidelines, this means that the outputs of generative AI systems (for example, deep fake content) are exempt from transparency obligations provided that they are not used outside of scientific research.

Commission Guidelines vs the Code of Practice

The European Commission (the Commission) has recently published two key documents on the

interpretation of Article 50 of the AI Act.

- On 20 July 2026, the Commission published guidelines on the implementation of the transparency obligations (“**Guidelines**”), and
- On 10 June 2026 the Commission published the final version of a Code of Practice on Transparency of AI-Generated Content (the “**Code of Practice**” or the “**Code**”).

These documents now form a framework to understand Article 50 of the AI Act, providing practical guidance to organisations on the implementation of the Article 50 transparency obligations.

**Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of the AI Act**

The Guidelines provide practical guidance on how the obligations are triggered, what various terms mean, and an introduction to the compliance obligations under Article 50 of the Act. The Guidelines are primarily structured around the four transparency obligations set out in Article 50 (as discussed further below).

The purpose of the Guidelines is to provide an interpretive framework, helping those subject to the transparency obligations to understand their legal scope and how they may be triggered in practice. While the Guidelines are not legally binding, in practice

organisations will likely need strong justifications as to why they did not comply with the Guidelines to avoid assertions of non-compliance.

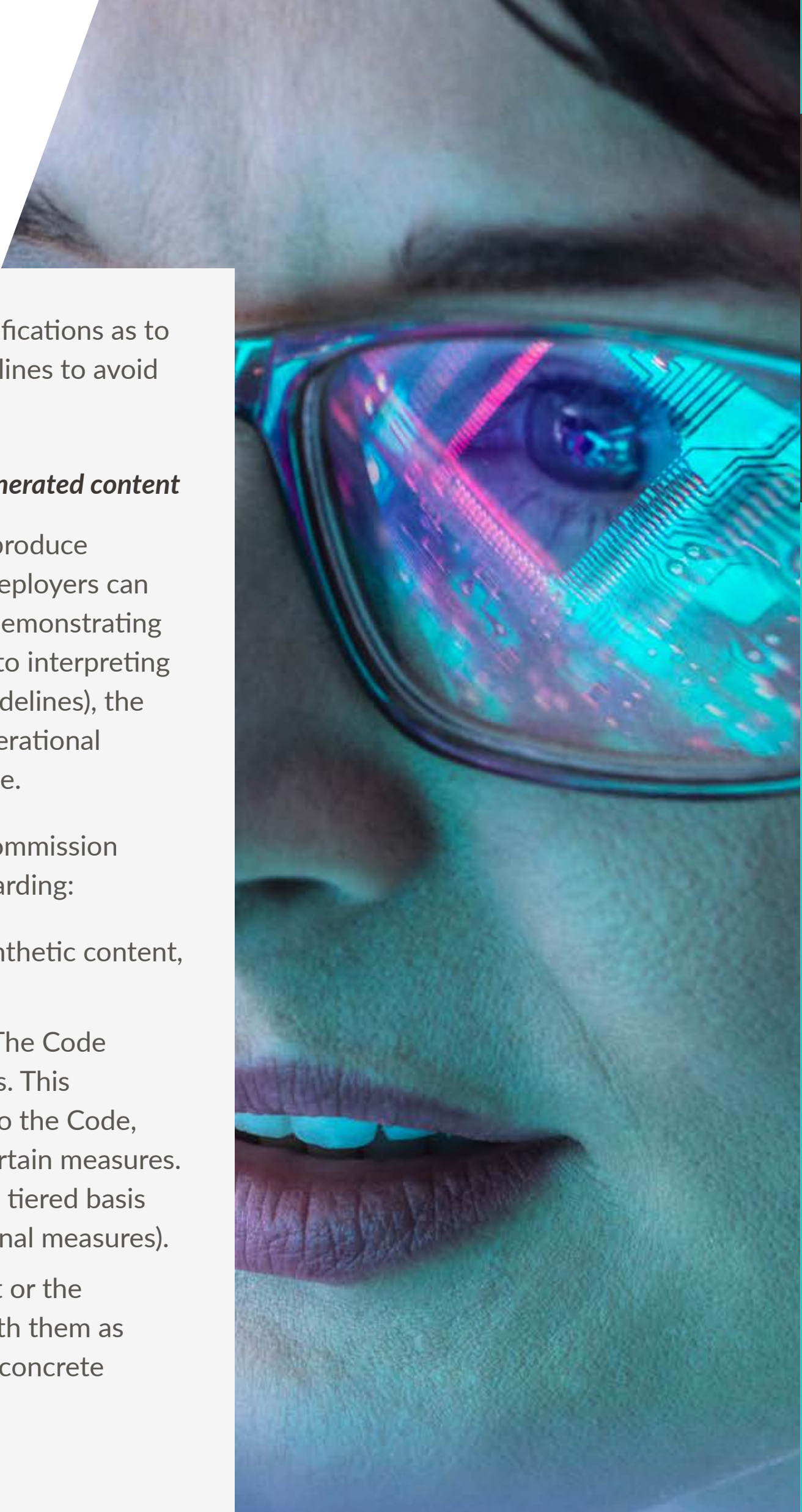
**Code of Practice on transparency of AI-generated content**

Under the AI Act, the Commission may produce “Codes of Practice” that providers and deployers can voluntarily participate in as a means of demonstrating compliance (“**Signatories**”). As opposed to interpreting the law (which is the function of the Guidelines), the Code of Practice provides a practical operational framework for demonstrating compliance.

The Code has been developed by the Commission specifically in relation to obligations regarding:

- The generation or manipulation of synthetic content, deep fakes, and
- Text published in the public interest. The Code operates on the basis of commitments. This means that by becoming a signatory to the Code, Signatories commit to carrying out certain measures. The commitments are structured on a tiered basis (mandatory, recommended, and optional measures).

The Code does not supersede the AI Act or the Commission’s Guidelines, but sits beneath them as a voluntary, operational tool setting out concrete measures that Signatories can adopt.





# Article 50(1): Requirement to build-in the transparency requirements into the AI system design

“*AI systems intended to interact directly with natural persons are designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system.*

(extract of Article 50(1))

APPLICABLE TO: PROVIDERS ONLY

Article 50(1) of the AI Act is aimed at AI systems that are intended to directly interact with individuals, requiring providers to design and develop their systems so that users are aware they are interacting with an AI system.

**Triggering the obligation**

To trigger the obligation, the AI system must be intended to directly interact with individuals. To demonstrate the necessary intention to interact, the AI system must be capable of responding to an input with contextual output. The Guidelines explain that “interaction” requires the possibility of a bidirectional exchange of information or actions. This interaction may take various forms including auditory, visual, or physical interaction. The interaction may also vary in duration and complexity. It can consist of a one-time exchange, such as a single prompt and reply, or a longer period involving multiple exchanges. By contrast, systems that only passively collect data, such as automated facial recognition access controls, cannot engage in an exchange with an individual. Similarly, AI systems that collect one-time feedback, such as spam filters, do not “interact” with individuals for these purposes.

The Guidelines state indirect or mediated human interaction where the person is exposed to AI outputs without directly dealing with the system itself do not fall within scope of this obligation. For example, if customer service representatives use AI assistance tools to help

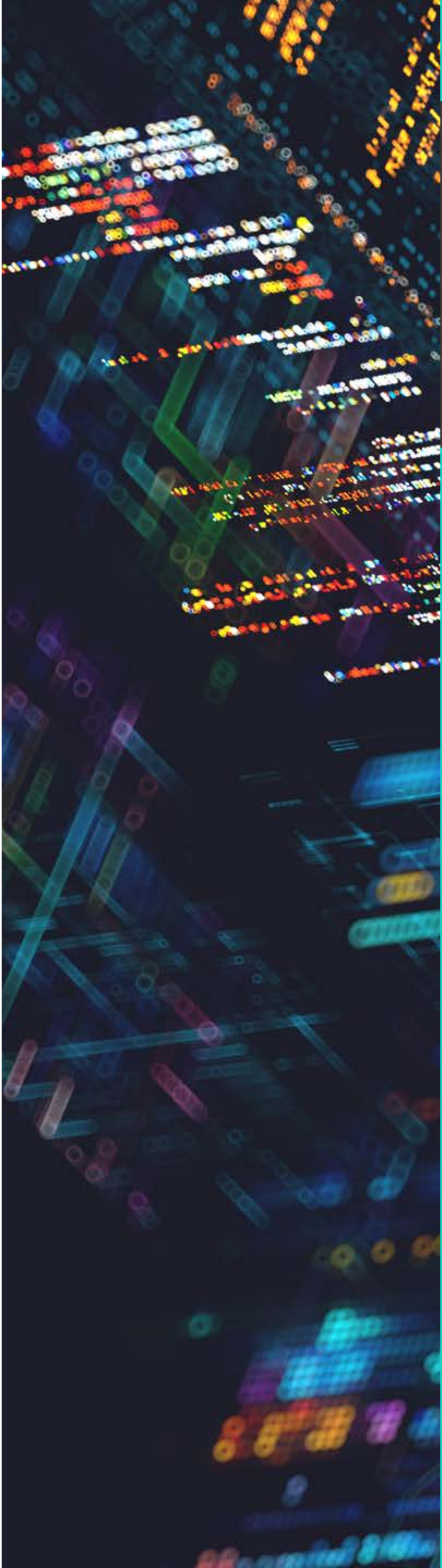
them communicate with individuals or if the AI output is not made available to the person by the AI system itself, but by another person disseminating the content.

Following some uncertainty regarding whether an expansive approach might be taken in the application of the Article 50(1) obligations, the Guidelines have provided useful clarification on which AI systems are deemed to be in scope.

- **Examples of AI systems in scope:** AI enabled voice assistance, chatbots, bots on social media networks, coding agents.
- **Examples of AI systems out of scope:** algorithmic recommender systems, spam filters, automated translation or transcription tools, backend decision support systems where the user only sees the AI output.

OBSERVATION

The emphasis within the Guidelines on an “interaction” i.e. a two way exchange of information being necessary to trigger this obligation is a welcome confirmation that will help organisations narrow the scope of the obligation.





## Complying with the obligation - informing users

The Guidelines emphasise that the obligation to inform users of their interaction with AI systems is ongoing and that natural persons must remain effectively informed throughout the lifecycle of the AI system.

**Format of disclosure:** The mechanism to inform the user should be embedded in the AI system. The disclosure technique should:

- Be provided in a clear and distinguishable manner, at the latest, at the time of the user's first interaction with the AI system, and
- Take into account the characteristics of the end users. Context is key here, with the Guidelines highlighting that the notification must be appropriate in the specific interaction environment. If the AI system is likely to interact with children, for example, then the notifications must be child-friendly and age-appropriate. The Guidelines suggest providers use standardised disclosure practices and labels to ensure effective and consistent notifications.

**Substance of the notification:** The notification should provide information on the artificial origin of the system that the individual is interacting with. The notification should ensure that individuals concerned are properly

notified in such a way that avoids risks of deception or manipulation before and throughout the interaction. The Guidelines give examples such as:

- **In writing:** A chatbot starts a conversation by mentioning that it is based on AI technology.
- **Visually:** An email generated by an AI agent features an AI label at the top.
- **Auditory:** A voice assistant says at the beginning of a session that it is powered by AI.

**Format of the notification:** A multimodal and accessible disclosure approach is considered best practice. The Guidelines suggest that different techniques could be combined to reinforce a user's understanding depending on the target audience and the size of the provider's organisation. Examples of possibly appropriate techniques (depending on circumstances) are given in the Guidance and include the following:

- **Textual disclosure:** Prominent plain-language labels or banners (e.g. you are interacting with an AI system) combined with badges or labels giving a visual indicator. It is suggested that accompanying disclosures should be placed close to the interaction interface (e.g. near the input/output field).
- **Auditory disclosure:** In a voice/telephony context explicit spoken statements (e.g. this is an AI powered

assistant) combined with periodic reminders with specific audio cues (e.g. tones) to support recognition.

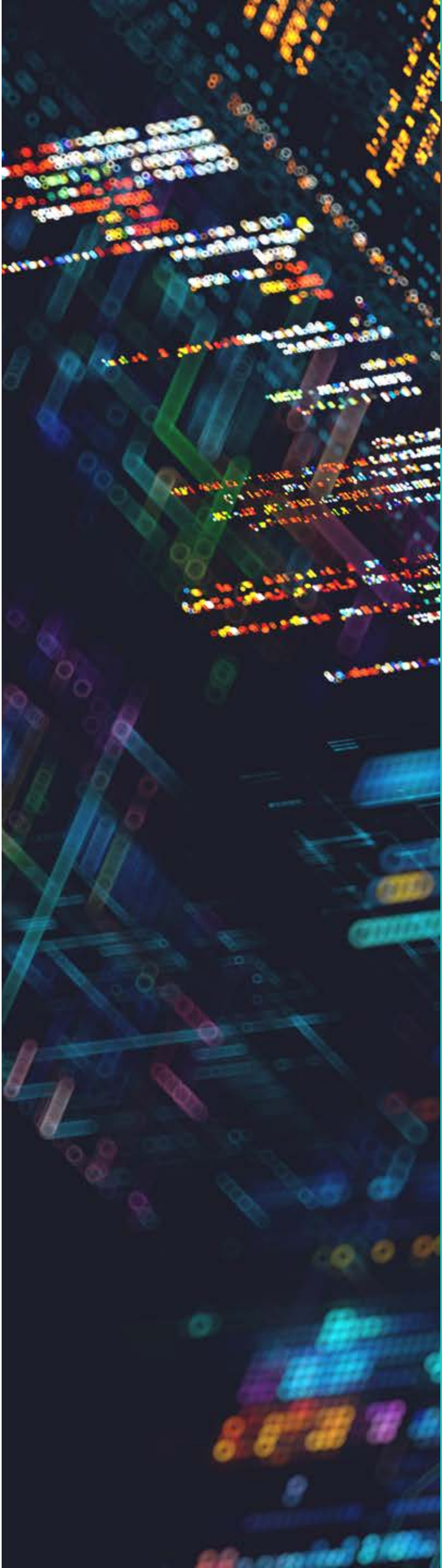
- **Visual/graphical cues:** Design of the AI system with watermarks, icons or recognisable frames to complement the textual disclosures are suggested.

The Guidelines also highlight that certain techniques, including disclosures contained in providers' terms and conditions and one-time disclosures, when used alone, will not be considered appropriate to effectively fulfil transparency obligations. The Guidelines also acknowledge the risk of overly intrusive disclosure techniques which decrease the effectiveness of the notification and lead to click fatigue and "banner blindness."

### OBSERVATION

The examples given in the Guidelines under both format and substance headings (as outlined above) are in some ways repetitive. The Guidelines highlight:

- the need to adapt to the user who is expected to be interacting with the system
- the use of plain/obvious language, and
- how a user might need to be reminded of the fact they are interacting with an AI system throughout their interaction.





## Exceptions

There are two limited exemptions to the Article 50(1) obligations.

### *Obviousness*

The first exception to having to inform users they are interacting with an AI system is where it would be obvious to the user that they are interacting with an AI system. The Guidelines note that this exception should be interpreted narrowly and be reserved to situations where *‘there is almost no doubt left’*.

- When considering what constitutes an “*obvious*” interaction with an AI system, the Guidelines note that the standard draws on established EU consumer protection law regarding the notion of “*average consumer*”. When considering the “*average*” individual, providers should consider
  - » their target audience, and
  - » how reasonably well-informed, observant, and circumspect an average member of that audience would be.
- Once the above two points are established, a provider should evaluate the obviousness by reference to the anticipated nature of the interaction and the characteristics of the target audience.

The Guidelines give some useful examples of scenarios in which the exception would and would not apply. These examples also demonstrate the importance of the target audience. For example, the Guidelines set out that the obviousness exception would not apply to AI chatbots embedded in online platforms or helpdesks. The exception would apply where an AI-powered code assistance chatbot is made available to professional developers, who by virtue of their expertise would satisfy the obviousness exception.

### *Law Enforcement*

Second, providers of interactive AI systems are exempt from the transparency obligations under Article 50(1) of the AI Act if they are authorised by law to detect, prevent, investigate or prosecute criminal offences.

The Guidelines provide an important caveat to this exception, noting that it does not apply where the AI system is made available to the public and the system offers a functionality to individuals to report criminal offences, such as police chatbots deployed on official police websites allowing individuals to report criminal offences.



# Article 50(2): Requirement to make the AI system output machine-marked and detectable

“AI systems, generating synthetic audio, image, video or text content, shall ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated.

(extract of Article 50(2))

**APPLICABLE TO: PROVIDERS ONLY**

Article 50(2) of the AI Act obliges providers to ensure that the synthetic output generated by an AI system is machine-marked and detectable.

## Triggering the obligation

Article 50(2) of the AI Act applies where an AI system is capable of generating synthetic content or manipulating content. Article 50(2) of the AI Act provides an exhaustive list of the types of in-scope content covered by the transparency obligation. This is quite broad and covers audio, image, video or text and includes multimodal content, which is content composed of a combination of these types of content (e.g. audio-visual content).

Interestingly, the Guidelines exempt source code and other outputs in the form of short sequences of numbers, symbols or letters from the obligations of Article 50(2). This is outlined in the Guidelines as meaning content written in a programming, scripting, markup query or configuration language intended to be interpreted, compiled or executed by the computing system.

Also not covered are AI systems that merely reproduce existing content or rearrange existing content. This may be the case, for example, with music playlists or recommender systems that only select or rank existing content without creating or manipulating anything.

The Guidelines explain that whether a person can perceive the content is a key factor in triggering this obligation. For example, if an AI agent takes steps on a user’s behalf and generates or alters content that a person can perceive, that output should be marked. However, where the agent carries out actions that do not generate content, such as internal processing steps, including chain-of-thought reasoning, or instructing a web browser to take an action, those actions do not qualify as synthetic content and are not subject to the marking and detection obligations.





## Complying with the obligation - marking and detection

Providers of AI systems in scope of Article 50(2) are obliged to comply with two distinct, but inherently connected obligations:

- marking AI generated or manipulated content, and
- ensuring the detectability of AI generated or manipulated content.

### Marking

Providers must ensure that outputs of the AI systems are marked in a machine-readable format. The Guidelines specify that technical solutions for marking and detection must be effective, interoperable, robust and reliable. The Code of Practice provides quality requirements in respect of these principles (e.g. Signatories are expected to measure the accuracy of the detection mechanisms). Signatories may also carry out research to contribute to the development of marking and detection mechanisms.

The Code includes specific requirements for marking content, by reference to four specific commitments.

- **Machine readable:** The marking must consist of at least one machine readable marking technique. Generally, a single layer of marking will not be considered sufficient but this should be assessed in the specific context of the AI system. Metadata must record information about the extent to which the content has been generated or manipulated by AI. An imperceptible watermark must be embedded into the content.
- **Non-removability:** Signatories must make best efforts to ensure that metadata and marking are preserved (i.e. Signatories will not intentionally change or remove them and will include an acceptable use policy and terms and conditions prohibiting such removal).
- **Provenance information:** Optionally, Signatories may consider incorporating richer information in the metadata to increase trustworthiness and integrity, for example information about the origin of the AI manipulated content, and
- **Perceptible watermarks:** Optionally, Signatories are encouraged to consider incorporating functionality for perceptible watermarks to assist deployers in their watermarking obligations.

### OBSERVATION

The Code itself recognises the practical reality that many AI systems are built on top of other AI models, the developers of which may be better placed to implement the Code's requirements. Indeed, the Code states that "*marking techniques may be implemented at different stages of the value chain (e.g. by the provider of the AI system itself or by an upstream model provider) and may also be provided by third parties*". The obligation and, ultimately, the responsibility for compliance, still rests with the provider of the AI system, not the provider of the underlying AI model.





### Detection

Under the Code, providers must also ensure that the outputs are detectable as artificially generated or manipulated. Signatories of the Code must provide a detection solution enabling users to verify whether content has been generated or manipulated by AI. The Code provides three types of detection solutions and encourages providers to consider specific points in connection with each:

- A public specification - providers should consider providing a reference implementation of the specification i.e. a worked example of the specification that would assist a third party in implementing a detection mechanism.
- Software - providers should consider ensuring that the software is operable on most operating systems including where possible mobile phones, or
- A cloud-based service accessible through an API for users in the EU.

The detection solution should be offered free of charge, although a Signatory with less than 1m monthly users may charge a reasonable fair and proportionate fee where the detection mechanism incurs substantial costs and

a single user exceeds a reasonable usage threshold for the detection mechanism. Signatories are required to ensure that their detection solution includes a detection mechanism for each marking technique implemented.

The Guidelines highlight that the provider is also obliged to make the means of detection available to the persons potentially exposed to the content and that the detection method must provide human-readable results. The Code notes that it must be managed in compliance with applicable legislation, such as data protection laws. Where a detection solution is withdrawn, a replacement must be implemented. Article 50(5) further mandates that the detection solution is clear and accessible to users.

Signatories to the Code may optionally consider providing a forensic detection mechanism capable of identifying AI-generated or AI-manipulated content, even where the marking has been removed. Additionally, Signatories may provide relevant information to deployers and other stakeholders to support their work in marking and implementing detection solutions.

Overall the solution(s) chosen to meet the requirements for marking and detection must be effective, reliable, robust and interoperable in the context of the overall state of the art.

The Code includes measures to support compliance. As such, Signatories are expected to:

- implement and continuously update an internal compliance process
- test marking and detection solutions before and after placing a generative AI system on the market. The Code contains an exemption for downstream providers, who may rely on upstream testing
- Train staff and external contractors involved in the development of AI systems, and
- cooperate with the relevant competent surveillance authorities under the AI Act.

Again, the Code notes that Signatories may rely on shared or third-party detection mechanisms, for instance provided by “*providers of AI models that apply markings at the model level*”. The Code encourages “*Signatories who are also providers of generative AI models*” to provide “*detection mechanism for the content generated or manipulated by their models prior to the model’s placement on the market*”



## Exceptions

### *Industrial application/B2B applications*

Where the AI output is strictly technical in nature and is intended to be perceived only by a limited pre-defined number of natural persons acting in their professional capacity within the organisation of the provider or deployer, such output will be exempt from the Article 50(2) obligations.

### *Assistance to standard editing or non-substantive alterations to input data*

The obligations under Article 50(2) of the AI Act do not apply if AI systems are used to check grammatical and technical errors, as long as the content is not materially amended. Similarly, AI systems intended to perform non-substantive alterations to the input data, or its semantics are excluded from the transparency obligation. The extent to which an alteration is considered substantial is context dependent. The Guidelines recommend following a case-by-case approach.

The Guidelines include numerous examples of this exception including grammar correction systems or

formatting technical compression or noise reduction or removal systems. Also excluded would be AI systems intended to impose minor cropping, colour adjustment or lighting adjustments to photos.

In contrast, the Guidelines give examples of AI systems that generate summaries or paraphrase text, remove objects from images, or create composite images, all of which fall within the scope of the obligation.

### *Law enforcement*

Article 50(2) lifts the requirement to mark AI generated content if authorised by law to detect, prevent, investigate or prosecute criminal offences.

### **TIMELINE FOR COMPLIANCE**

The recent passage of the AI Omnibus extends the timeline for compliance with the obligations connected to the generation and/or manipulation of synthetic content (Article 50(2) of the Act) for AI systems already on the market on 2 August 2026 to 2 December 2026.



# Article 50(3): Requirement to inform individuals impacted by emotion recognition or biometric categorisation system

“*Emotion recognition system or a biometric categorisation system shall inform the natural persons exposed thereto of the operation of the system.*  
(extract of Article 50(3))

APPLICABLE TO: DEPLOYERS ONLY

In respect of those uses that are not prohibited, Article 50(3) of the AI Act requires deployers of emotion recognition or biometric categorisation systems to inform natural persons that they have been exposed to the operation of such AI systems. Deployers are required to process biometric data in accordance with applicable legislation, including data protection legislation. The Article 50(3) transparency obligation is in addition to the high-risk obligations that may apply to emotion recognition and biometric categorisation systems which are separate and also may apply.

## Triggering the obligation

Article 50(3) applies specifically to:

- **Emotion recognition systems:** these are systems designed to identify or infer a person’s emotions or intentions from biometric data (e.g. facial expressions, voice, physiological signals, but not physical states such as pain or fatigue). The Guidelines note that all emotion recognition systems are also classified as high-risk under the AI Act and therefore will need to also comply with the high-risk obligations under the AI Act.
- **Biometric categorisation systems:** these are systems that classify individuals based on specific characteristics (e.g. age group, gender, behavioural traits) based on biometric data . It applies regardless of whether the AI system falls in the scope of high-risk AI systems under the AI Act. The definition of biometric categorisation systems excludes biometric categorisation systems that are used as an “ancillary” component to a main “commercial service”. To be able to avail of this exception such use must

be strictly necessary to meet objective technical reasons. This would occur where the biometric categorisation system could not be used without the main commercial service. For example, a system that categorises people’s facial and body features to allow customers on an online retail marketplace to preview how a product would appear on them before they buy it. The main commercial service is the sale of products online, with the filtering system being ancillary to that activity.

The obligation does not apply where permitted by law to the detection, prevention or investigation of criminal offences, subject to implementing appropriate safeguards. This exception is targeted at official law enforcement use.



## Complying with the obligation - informing those exposed

The core obligation for deployers under Article 50(3) is to inform individuals that they are exposed to an emotion recognition or biometric categorisation system. Deployers may inform individuals in advance, but they must do so at the latest by the time of first exposure to the AI system. This obligation applies equally where the exposure does not occur in real-time.

The Act provides significant flexibility as to both the means and content of the information provided. Deployers are not required to explain the reasons for the system's operation. The means of disclosure are context-dependent and may take the form of written notice, oral communication, or standardised icons. By way of example, a gallery might display a notice at the entrance to each exhibition room, while a computer game might include a pop-up message before gameplay begins. The Guidelines recommend that deployers also consider the broader context, including the location and the nature of the relationship with the individual receiving the notice, for instance, any disabilities that may affect how the information is received.



# Article 50(4): Requirement to disclose that a deep fake or text published in the public interest has been artificially generated or manipulated

“AI system that generates or manipulates image, audio or video content constituting a deep fake, shall disclose that the content has been artificially generated or manipulated.”  
(extract of Article 50(4))

“AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest shall disclose that the text has been artificially generated or manipulated.”  
(extract of Article 50(4))

APPLICABLE TO: DEPLOYERS ONLY

Article 50(4) requires the deployers of AI systems to disclose to individuals that a deep fake or text published in a public interest has been AI-generated or manipulated.

## Triggering the obligation

Article 50(4) of the AI Act imposes two main obligations on deployers of AI systems:

- **Deep-fakes:** To disclose that content such as images, audio or video constituting a deep fake has been artificially generated or manipulated; and
- **AI-generated or manipulated text in the public interest:** To disclose that text published for the purpose of informing the public on matters of public interest has been artificially generated or manipulated.

## Deep fakes - obligation trigger

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.” The Guidelines explain that, to constitute a deep fake, four key criteria must all be satisfied.

**1. Resemblance:** The AI-generated or manipulated content must resemble an existing subject, meaning there must be a high level of similarity between the deep fake and its real-world counterpart. This requires

a case-by-case, objective assessment.

- 2. Existing subject:** The content must depict a subject that already exists in the real world. This criterion works in tandem with the resemblance requirement. The AI output must not only look realistic but must be modelled on something or someone that actually exists.
- 3. Persons, objects, places, entities, or events:** The definition enumerates the specific categories of subject matter that may be depicted. The Guidelines clarify that these should be understood as follows:
- » persons: realistic human beings, including digital replicas
  - » objects: realistic inanimate items, such as buildings or artwork
  - » places: realistic locations
  - » entities: realistic non-human but animate beings, such as animals, and
  - » events: realistic scenes or situations
- 4. False appearance of authenticity:** The essential characteristic of a deep fake is its ability to deceive. The content must falsely appear to be genuine, that is, a viewer must be liable to perceive it as authentic and truthful when it is not.





One example of where organisations will need to consider the deep fake obligation are marketing departments who may look to generate realistic images for the promotion of the business.

Importantly, this assessment is audience-specific. The Guidelines make clear that the perception of the specific audience the deep fake may reach must be considered on a case-by-case basis, rather than by reference to an “average person” standard. It follows that not every AI manipulation of content will constitute a deep fake, what matters is the likely impact on the particular audience in question.

### Complying with the obligation - disclosing the deep fake

AI-generated content must be “clearly and distinguishably” labelled in a visible manner that is accessible to individuals without the need for any additional measures.

The AI Act does not prescribe the practicalities of these obligations in detail, so the Code of Practice provides helpful guidance. Where visual disclosure is possible, Signatories are required to:

- use the “AI” acronym in English, unless otherwise prescribed by national law
- place the label vertically, with its size proportionate to the content, to ensure it does not mislead
- ensure the acronym remains readable, even where stylistic variations are used

- consider implementing a second, interactive layer, recommended by the Code of Practice to achieve greater transparency, that explains whether and how the content has been manipulated or generated by AI.

The EU has also provided suggested icons that can be used, reproduced below.

| BASE ICON                                                                           | FULLY AI GENERATED                                                                  | FULLY AI GENERATED                                                                  |
|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
|  |  |  |

Deployers have two choices: they can either use the EU icon or apply their own icon or equivalent label. Both options must be accompanied by a description, with the EU icon carrying standardised wording. The Code also sets out placement principles, requiring deployers to consider:

- the duration for which the label remains visible
- the need for the label to be embedded within the content, and
- as previously noted, the requirement that it be clearly distinguishable.

Where visual disclosure is not possible, the Signatory should consider audio disclosure, which, in any case, must be accompanied by visual materials.

### Exceptions

The Guidelines note that the AI Act envisages transparency obligations for deep fakes being implemented in a way that does not detract from evidently artistic, creative, satirical, or fictional work. However, in our view, this exception is likely to be narrowly interpreted.

### Relation between Art. 50(4) of the AI Act and Art. 35(1)(k) of the DSA

The Digital Services Act (“**DSA**”) requires providers of very large online platform providers (“**VLOPs**”) and very large online search engines (“**VLOSEs**”) to consider marking “*generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful*”.

As the Guidelines note there are two material differences.

- The obligations are aimed at different users. Under the DSA the obligation is aimed at the provider of the service (the VLOP/VLOSE), whereas under the AI Act, the obligation is aimed at a user who may be the ‘recipient of the service’ under the DSA.
- The obligation under the DSA covers a broader range of content which may appear false. The AI Act applies only to content that is created by an AI system.



## Texts published in the public interest - obligation trigger

Deployers have a transparency obligation in relation to AI-generated or manipulated text that informs the public on matters of public interest. In order to fall within the scope of this provision, the text must satisfy three conditions:

- **Published:** the text must be accessible to a fairly large, indeterminate number of unrelated potential readers, whether simultaneously or successively, regardless of whether access is provided against payment. This threshold is drawn from the Guidelines.
- **Informing the public:** the aim of the text must be to communicate knowledge, opinions, or facts to the public.
- **On matters of public interest:** the subject matter must be relevant to society at large, covering issues that merit public debate or scrutiny.

The aim of this obligation is to apply to organisations that use AI tools to generate public interest text.

## Complying with the obligation - disclosing that the text is AI-generated/manipulated

When it comes to public interest text, the core obligation is transparency. Any text that has been artificially generated or manipulated must be disclosed as such. As with deep fakes, deployers are required to mark their content “*clearly and distinguishably*” in a manner that is visible and accessible to natural persons, without the need for any additional or assistive technology.

The Code imposes a number of specific obligations on Signatories, including the following:

- **Internal compliance:** Proportionate to their size. Signatories should maintain internal compliance processes and documentation that specifies how they implement the disclosure obligations using the icon or equivalent label. Signatories are also encouraged to publish information on their disclosure solution.
- **Visual disclosure:** Where visual disclosure is possible, the Code sets out design specifications. These mirror those already discussed above in the context of deep fakes.
- **Awareness and training:** Signatories must make efforts proportionate to their size to ensure that personnel, employees, and external contractors are aware of the disclosure obligations. This includes providing training or equivalent guidance on when and how disclosure is required.



### Exceptions

Transparency obligations under Article 50(4) for text published in the public interest do not apply where the AI-generated or manipulated content:

- has undergone human review or editorial control, and
- is published by a company or individual that holds responsibility for the publication.

The Code provides that human review and editorial control can be established in two ways.

- Signatories who are media service providers within the meaning of Art 2(2) of the European Media Freedom Act (Regulation 2024/1803) and already subject to editorial standards, and to regulatory or co-regulator or self regulatory frameworks can use their existing editorial procedures and standards to rely on the exception.

- All other Signatories must establish internal procedures that, at a minimum, identify an individual or entity with editorial responsibility and provide an overview of the organisational measures in place to ensure adequate human review and editorial control. These procedures effectively require an organisation to name who is responsible for the output and supervision of the AI generated text and put in place processes to operationalise that responsibility. This aims to reduce the risks associated with using AI systems to publish text that it is in the public interest. Editorial responsibility must be assumed before publication, and the contact details of the responsible person must be made publicly available.



Key contacts



**John Cahir**  
*Partner*  
+353 1 649 2943  
jcahir@algoodbody.com



**Chris Bollard**  
*Partner*  
+353 1 649 2328  
cbollard@algoodbody.com



**Stephen King**  
*Partner*  
+353 1 649 2243  
sking@algoodbody.com



**Aideen Burke**  
*Partner*  
+353 1 649 2627  
aiburke@algoodbody.com



**Andrew Sheridan**  
*Partner*  
+353 1 649 2766  
asheridan@algoodbody.com



**James Green**  
*Associate*  
+353 1 649 2702  
jgreen@algoodbody.com



**Katie O'Neill**  
*Associate*  
+353 1 649 2529  
katoneill@algoodbody.com

Article contributors:

Kate Kilcommins, Aoibhin Collins, Elisa Crespo Miguelez, Justyna Palczynska and Molly McHugh

Resources



[The EU AI Act](#)

READ MORE



[The Digital Omnibus on AI](#)

READ MORE



[European Commission guidelines on transparency obligations](#)

READ MORE



[Code of Practice on Transparency of AI-Generated Content,](#)

READ MORE