

ANNEX – SELECTED FAQs

(Source: European Commission)

1. When must deployers of AI-based emotion recognition or biometric categorisation systems inform natural persons exposed to those systems?

According to Article 50(3) of the AI Act, deployers of emotion recognition systems and biometric categorisation systems must inform natural persons who are exposed to those systems of their operation to protect their privacy. This does not entail providing information regarding e.g. the purpose of the system's operation.

The obligation applies regardless of whether the persons are exposed to such systems in real-time or they are operated ex-post.

2. What is a 'deepfake' and when do they need to be labelled?

Deepfakes are defined in Article 3(60) of 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.

Three cumulative criteria need to be met for content to constitute a 'deepfake':

1. Resemblance: a high level of similarity between the deepfake content and the simulated subject.
2. Existing: simulated persons, objects, places, entities or events need to resemble someone or something that exists, can plausibly exist or could have plausibly existed in reality.
3. False appearance to be authentic or truthful: this relates to the essential characteristic of deepfake content and its capacity to potentially deceive or mislead a person regarding the content's authenticity or truthfulness

For the assessment of the last criterion above, the Guidelines on Transparency of AI-Generated Content allow stakeholders to take into account:

- the level of resemblance,
- the potential substantive message of the content
- the intended and foreseeable deployment contexts,
- the intended and reasonably foreseeable audience composition and their expectations.

This enables deployers to duly consider the specific context in which they are operating and the audience to which their content will be exposed. If the intended audience within a specific deployment context does not expect the content to be authentic or truthful, the AI-generated or manipulated content may not falsely appear to be authentic or truthful. For example, the AI-generation or manipulation of background scenes, special effects, or technical pre- and post-processing as part

of standard movie production processes are not likely to make content falsely appear to the audience to be authentic or truthful.

Deployers must disclose deepfake content to a natural person upon first exposure at the latest. This disclosure should happen in a clear and distinguishable manner. It should be understandable and perceivable by natural persons (e.g. with visible or audible labels), without need for any specific technical tools or performing dedicated actions. Therefore, deployers cannot simply rely on the machine-readable marking embedded in the content by the provider under Article 50(2) of the AI Act to fulfil their disclosure obligation.

Some deepfakes are part of evidently artistic, creative, satirical, fictional or analogous works or programmes. For these, the transparency obligation is limited to the disclosure of the deepfake content in an appropriate manner that does not hamper the display or enjoyment of the work.

4. What AI-generated or manipulated text must deployers clearly label?

According to Article 50(4) of the AI Act, deployers of generative AI systems must clearly label AI-generated or manipulated text published with the purpose of informing the public on matters of public interest. For text to fall within this obligation, 3 criteria must be fulfilled.

The text needs to be:

- **Published**
- **Informative to the public**
- **On matters of public interest**, such as:
 - politics and democratic processes,
 - public administration and services,
 - administration of justice and law enforcement,
 - fundamental rights,
 - public security,
 - public health,
 - environmental protection,
 - consumer safety and any economic,
 - financial, political, scientific, or cultural developments that may
 - be relevant subject of public debate.

5. What constitutes the human review or editorial control/responsibility required to qualify for an exemption from the labelling obligation?

Published text that has undergone human review or editorial control – does not need to be labelled.

Human review refers to the deliberate examination of the substance of the content by one or more natural persons possessing relevant knowledge and professional judgement pertaining to the subject matter under scrutiny (e.g. academic peer review or professional validation chains). Editorial control refers to the control exercised in practice by a responsible editorial entity (e.g. an editor-in-chief) over the content having the authority to approve, alter or reject the substance of the

text based on substantive grounds (incl. factchecking of information and ensuring the trustworthiness of sources).

Superficial, solely formal, or procedural checks (e.g. spell-checking or grammatical correction) are not considered to be human review or editorial control. Editorial responsibility means that a person must hold the ultimate legal responsibility over the publication of the content, including the human review or editorial control.

6. How can providers and deployers demonstrate compliance and what are the effects of the Code of Practice on transparency of AI-generated content?

The Code of Practice on Transparency of AI-generated content provides a voluntary practical tool to help providers and deployers of generative AI systems demonstrate compliance with the marking and labelling obligations under Article 50(2), (4) and (5) of the AI Act. The code has been assessed as adequate by both the Commission and the AI Board. It allows signatories to benefit from legal certainty, predictability, and trust as they may rely on an approved code that they have signed, regardless of their place of establishment, operation or competent supervisory authority.

Providers and deployers of generative AI systems that decide not to adhere to the code will have to demonstrate compliance through alternative adequate means. They may be subject to more requests for information, since there is less transparency on how they comply with the transparency obligations of Article 50(2), (4) and (5) of the AI Act. For the other transparency obligations (e.g. those laid down in Article 50(1) and (3) of the AI Act), providers and deployers can determine adequate compliance measures themselves, while taking into account the Guidelines on Transparency of AI-Generated Content.

7. How does the Code of Practice on transparent generative AI systems interact with the recent Code of Practice on General-Purpose AI?

The transparency obligations laid down in Article 50 of the AI Act are complementary to the transparency rules applicable to general-purpose AI models laid down in Articles 53 and 55 of the AI Act. The means to comply with the latter have been further detailed in the Code of Practice for General-Purpose AI (GPAI) (<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>) models and the Commission template for the summary of the content used for model training.

Notably, the GPAI Code of Practice focuses on GPAI models and documentation and information to be provided to the AI Office, national competent authorities and downstream providers, and transparency of the input training data. The Code of Practice on transparency of AI-generated content addresses marking techniques and transparency of the AI generated or manipulated outputs towards the persons exposed to them.

The Code of Practice on transparency of AI-generated content targets transparency obligations at the system level, including, but not limited to, GPAI systems. Transparency techniques which can be implemented by providers of

GPAI models to facilitate compliance by downstream AI system providers with the transparency obligations of Article 50 of the AI Act will also be considered in the context of this code.

8. When does Article 50 of the AI Act start to apply and is there a grace period?

Article 50 of the AI Act (<https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-50>) applies as from 2 August 2026. From that date onwards, providers and deployers of AI systems must comply with the transparency obligations laid down in that provision. A limited grace period is envisaged only for AI systems placed on the market before 2 August 2026 and only as regards the marking and detection obligation for AI-generated content (Article 50(2) of the AI Act). Providers of such systems must comply with those obligations only as from 2 December 2026.

Content generated prior to 2 August 2026 does not need to be labelled retroactively. Nonetheless, the Commission encourages relevant deployers to do so, where possible, since it contributes to the goals pursued by Article 50 of the AI Act.

9. Who will oversee the enforcement of Article 50 of the AI Act and have the power to impose fines?

Compliance with the rules will mainly be enforced by national competent market surveillance authorities. The AI Office has a limited role in monitoring and enforcement, since it is only competent for AI systems that are built on general-purpose AI models, if the same entity provides the system and the model, or if the AI system is integrated into a very large online search engine or very large online platform designated under the Digital Services Act (<https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops>). The European Data Protection Supervisor (https://www.edps.europa.eu/_en) will enforce the rules vis-à-vis AI systems used by the EU institutions, bodies and agencies.

Fines can reach up to 15 million euros or 3% of total worldwide turnover for the preceding financial year, while proportionality can be taken into account in the case of small and medium sized enterprises (SMEs) and small mid-cap companies (SMCs).

EXAMPLES

Examples of deep fakes under Article 3(60) AI Act

- AI-manipulated image of a scene involving two real professional footballers in front of a building resembling a football stadium.
- AI-generated audio involving voice cloning of a newspaper podcast's regular presenters and a guest interlocutor discussing some news events.
- AI-generated video of an individual resembling a politician, holding a speech in front of an audience.
- AI-generated video featuring an AI-generated depiction of a celebrity influencer in an advertising or promotional context.
- AI-generated video featuring a realistic synthetic avatar of a company CEO congratulating employees with their work and the corporate results of the past year.
- An AI-generated image of a product in advertisement or packaging that can affect the audience's perception and mislead as to the actual product appearance, characteristics or use (e.g. making the product appear not identical to the real product, more appealing or with improved quality than in real life).

Examples that do not constitute deep fakes under Article 3(60) AI Act

- AI-generated image of a scene involving a sphinx flying over the Eiffel Tower.
- AI-generated video of mice arguing in human language over the best type of cheese as part of an advertisement campaign for a cheese manufacturer.
- AI voice replication for fictional characters (e.g., for audiobooks, games or animation) when there is no deception as to the identity of the narrators.
- AI-manipulated radio broadcast adjusting technical audio parameters (e.g. normalising volume levels, noise reduction, audio compression) without altering the actual words spoken by speakers or their way of speaking.
- AI-generated cartoon of a pre-existing image depicting a historical event.
- Video of a presenter of a scientific TV-programme who walks across a mountainous region while addressing the retreat of glaciers which is illustrated by an AI-manipulated background animation that shows how the glaciers retreated from their historical levels to their current levels.
- AI-generated fictitious environments (e.g. fictitious forests and castles) in video games.
- Movies where real existing actors (not manipulated by AI) are playing against an AI-generated background (e.g. depicting a fictitious stretch of prairie, historical buildings of ancient cities etc.).
- A real product (e.g., a car) shown in an advertisement against an AI-generated background and surrounding environment as long as the ad is not likely to mislead the audience about the product's actual representation and its characteristics and use.

Examples of deep fakes that are creative, satirical, fictional or analogous works

- Artistic/fictional work: movies (or movie trailers) featuring AI-generated or manipulated deep fakes of de-aged, existing actors or digital replicas of dead actors or other persons, available in a cinema or on a streaming platform.
- Artistic/creative work: AI-generated music in any kind of genre resembling the individual style of existing artists.
- Satirical/fictional work: AI-manipulated image of an existing politician placing him in a scene clearly meant to criticise in a humorous manner certain policy decisions taken or supported by that person.
- Analogous creative/fictional work: AI-generated gaming imagery involving deep fake simulations of real, existing persons.

Examples of deep fakes that do not constitute artistic, creative, satirical, fictional or analogous work:

- AI-manipulated video in the style of a teleshopping channel involving deep fake simulation of humans advertising a product in an AI-generated scene depicting the functional use of the product by the simulated consumers with the aim of persuading viewers to buy the product.
- AI-generated image of celebrities implying their involvement in activities that never happened, lacking any fictional, satirical or analogous purpose.
- AI-manipulated video featuring a realistic synthetic influencer testing out a sponsored real product, focused solely on displaying the product functionalities.
- AI-generated video depicting realistic, holocaust scenes being shared on publicly available social media platforms.

Examples of text published with the purpose of informing the public on matters of public interest under Article 50(4) AI Act which must be labelled as artificially generated or manipulated unless subject to human review/editorial control (see Section 6.2.3.):

- AI-generated summary of a human-authored article on a newspaper's website discussing a recent decision by a town council.
- AI-manipulated parts of a lifestyle-website article comparing the effects of various diets on a particular disease in middle-aged women.
- AI-manipulated corporate reports published on a listed company's website containing investor information.
- AI-generated message on a meteorological institute's social media profile warning citizens about stormy weather and related precautionary measures.

Examples of text which is not published with the purpose of informing the public on matters of public interest falling outside the scope of Article 50(4) AI Act:

- AI-generated fantasy novels.
- AI-manipulated text that is part of a company's advertisement or product descriptions (not including any claims related to e.g. health, consumer safety or sustainability).
- News summary generated by a chatbot that is only available to the user that prompted the chatbot.
- AI-manipulated text by a consultant for a client advice regarding measures to be taken for regulatory compliance with applicable legislation.

Examples of standard editing and minor alterations benefitting from the exception under Article 50(2) AI Act:

- Grammar correction and spellchecking, linguistic and minor stylistic polishing that do not change the substance, meaning, style or messaging of text, AI-generated translations of text;
- Formatting, format conversions, technical compression, noise reduction or removal for enhanced clarity without changing the meaning or the substance of the content;
- Minor cropping, minor colour adjustments or corrections, lightening or darkening, sharpening for enhanced clarity or other standard editing and technical corrections;
- Removal of dust spots caused by a dirty lens or sensor, removal of red-eye caused by flash photography; deleting and obscuring backgrounds that are visible in the original file, pixelation or blurring of faces;
- Rescaling of a video clip, dynamic range compression and equalisation; limited video stabilisation;
- Standard adjustments to colour and contrast, minor adjustments to playback speed, minor corrections to level the horizon of an image, applying pixel filters to amplify certain parts of an image, or applying colour maps to grayscale images, edge image completion, converting a black & white to colour image or video and vice versus;

- Pixel filling for aspect ratio adaptation or enhanced clarity, automatic transition clips, other non-substantive edits for cosmetic and technical purposes.
- Medical image technical processing or reconstruction within medical devices (e.g. annotations, boxes, contours, circles, heatmaps);
- Transcriptions of conversations; AI-generated content that only transforms authentic human input through assistive technologies allowing persons with disabilities to communicate (e.g., augmentative and alternative communication (AAC) or customized neural voices (CNV)) since they do not alter semantically the meaning of the content.

Examples of semantic changes that require marking under Article 50(2) AI Act:

- AI-generated summaries of text; paraphrasing or rewriting text that changes style, structure and meaning beyond mere grammatical and minor stylistic correction;
- Removal, replacement or insertion of objects or persons in existing images and videos that changes meaning and substance of the content; face replacement or substantial facial modification,
- Synthesis of realistic speech in a specific person's voice, or generation of realistic video depicting events that did not occur; altering the body shape or the skin colour of a person; extreme lightening, darkening, colour and contrast changes that change the meaning, intent and messaging of the content;
- Creation of composite images or video clips that modifies the representation of persons, objects, events or facts, any other substantial alteration of the content.