

EU AI transparency rules: what they mean for your advertising (and why UK brands should care too)

Overview:

The EU AI Act's new transparency rules mean many ads using AI must now need clear "deepfake" labelling, including UK brands whose campaigns reach EU audiences online. This article unpacks when the rules arise, who is responsible, what counts as a deepfake, and how to label ads correctly.

Full article:

EU AI transparency rules: what they mean for your advertising (and why UK brands should care too)

In brief

- Since 2 August 2026, the EU AI Act has introduced new requirements to make the use of AI more transparent, with advertising a key area directly affected.
- This article explores what this means in practice for brands, marketing teams, agencies and their in-house legal advisers: when do these rules apply to your advertising, what constitutes a 'deepfake' under the rules, what do you actually need to do if the rules do apply, how do you label ads, and does the answer change depending on the type of media?
- With the industry currently grappling with these questions and more, this article aims to bring clarity to what is a significant shift for both the industry and for compliance.
- Getting it wrong carries real financial risk, so AI labelling for in-scope ads needs to become a standard part of the creative process from the outset, in the same way that other on-screen disclosures, such as comparative advertising supers or safer gambling messaging, have already become part of everyday advertising.

Guidance and Codes of Practice

Throughout this article, we refer to the European Commission's [Guidelines on the implementation of the transparency obligations](#) (Guidelines) as well as the [Code of Practice](#) which supports compliance with the EU AI Act's transparency requirements.

When are the EU's transparency rules applicable?

There are two practical questions to work through: first, does this apply to you at all given where your ad is shown, and second, whose responsibility is it to comply. We take each in turn below.

1. Do the rules apply to your advertising?

Whether the rules apply to your advertising is a question of geography, but not simply whether an ad happens to reach the EU. An ad falls within scope if you foresee, direct or authorise it reaching an audience in the EU, wherever the brand, agency or servers are based, including where it is simply posted on the globally accessible internet. This test is likely to be easy to satisfy and hard to avoid, given how advertising works in practice. Is it realistic to run a labelled version of an ad for EU audiences and an unlabelled version for the UK, and police which one goes where? And, even if you wanted to, most digital, social and online video placements cannot be reliably geo-blocked, so an ad put online for a UK audience can very easily still reach viewers in the EU.

Given that most global and European brands run substantially the same creative across the UK and the EU, and the UK has no equivalent of Article 50 requiring a separately labelled version (currently, a consultation is expected imminently), the safest and practical approach is to treat EU-standard labelling as the default for any ad that uses AI in a way that could require it, rather than trying to maintain two versions or rely on geography to avoid the obligation. That said, a campaign that is genuinely and deliberately targeted only at the UK, and that reaches the EU only incidentally and outside the advertiser's control, for instance if someone privately forwards it, would likely remain defensible as outside scope. It is only where EU exposure is foreseeable that the safer approach above should be followed.

It is also worth noting that the obligation is not retrospective for image, audio and video deepfakes: content generated or manipulated before 2 August 2026 does not need to be labelled, even if it continues to run or be republished after that date. However, the position is different for AI-generated or manipulated text on matters of public interest, where the relevant date is publication rather than generation, so text produced before 2 August but first published on or after that date will need to be labelled, unless it benefits from the editorial control exception. Even where labelling isn't required, businesses are encouraged to label older content where they can, though they aren't expected to make disproportionate efforts, such as auditing content databases, to do so.

Once you've concluded the rules are likely to apply, the next question is whose responsibility it is to comply.

2. Who does this apply to: are you a deployer or a provider?

The EU AI Act distinguishes between 'deployers' and 'providers'. In an advertising context, brands and agencies will almost always be deployers: the party deciding whether and how an AI tool is used to create or manipulate an ad. You will only also be a provider in the much narrower case where you have built and are using your own in-house AI system to generate that content, rather than using a third-party tool such as an image, voice or video generator. We mention providers here because some agencies do build their own generative AI tools, and, where that happens, they will take on separate obligations in that capacity, in addition to obligations as a deployer. Most brands and agencies, however, will only need to focus on the deployer obligations covered in this article.

Within that deployer category, it also matters whether it is the brand or their agency that is treated as the deployer. The Guidelines suggest that a brand commissioning an ad will generally be treated as the deployer if it decides whether and how AI is used in the ad. In practice, a brand will usually have final sign-off over how the ad looks and is produced, so it should expect to be treated as the deployer even where an agency does the day-to-day production work. The exception is where a brand genuinely leaves those decisions to the agency and does not exercise control over whether or how AI is used, in which case the agency itself would be the deployer, though this would be an unusual arrangement in most client relationships.

That said, a brand and its agency can each independently be a deployer on the same campaign. For example, a brand might set the overall rules for how AI can be used and sign off the final creative, while the agency retains day-to-day control over which specific AI tools are used and how. Both of those roles can count as exercising authority over the AI system, so responsibility can genuinely be shared rather than resting solely with whoever technically pressed the button. The practical upshot is that this isn't something either party can ignore or pass on responsibility to the other. The safest approach is to build AI labelling into the conversation from the very start of a campaign: when agreeing the brief and scope, and agreeing which elements will use AI, agree at the same time whether a label will be needed, what it will say, and where it will go, so that labelling becomes a standard part of the creative sign-off process rather than an afterthought raised only once the campaign is ready to go live.

An advertising agency could also hold both roles (provider and deployer) at once. For example, an agency that builds and uses its own in-house generative AI tool which is used to create a deepfake ad would be both the provider (subject to the separate marking and detection duty in Article 50(2) of the EU AI Act) and the deployer (subject to the Article 50(4) of the EU AI Act labelling duty).

You can read more about the other transparency obligations in our article: [What the EU's final AI transparency guidelines mean for providers and deployers](#).

How does the obligation play out in practice, by media and format?

As discussed above, whether an ad is likely to reach an EU audience is what determines whether the labelling obligation bites. In practice, this plays out differently depending on the media used. Broadcast media, such as TV and radio, can often genuinely be confined to a single market, whereas online media, such as websites, social media and online video, is much harder to keep within one country once it's live.

Take an AI-manipulated advert featuring a synthetic spokesperson that runs on UK television, YouTube, social media and the brand's website. The television broadcast might realistically stay within the UK, but the same ad on YouTube, social media or the brand's website is a different story. EU users can access that content directly, so it becomes far harder to argue reaching an EU audience wasn't foreseeable.

Social media deserves a specific mention here, given how much advertising now runs through it. Many platforms, including Instagram and TikTok, already offer their own built-in AI labelling tools, whether applied automatically or added voluntarily by the person posting. The difficulty is that these platform labels won't necessarily satisfy the EU AI Act labelling obligation on their own: for example, if the platform's label isn't itself clear and prominent to the viewer, for example because it's only embedded as invisible metadata, it won't count as compliant disclosure. Even where it is visible, the advertiser and agency remain responsible for checking that it actually meets the requirements. The simplest way around this is not to rely on the platform at all; if labelling is built into the ad itself as part of the production process, as suggested above, it travels with the content regardless of which platform it ends up on.

The upshot: a broadcast-only placement may be the one scenario where the geography point genuinely works in your favour, but it rarely helps once any part of the same campaign goes online, which is true of almost every campaign today.

What type of AI generated content triggers the labelling requirement?

The EU AI Act uses the word "deepfake" to describe what needs labelling, but this extends to AI-generated use that goes beyond the fake celebrity videos most people picture when they hear that word. Under the EU AI Act, a deepfake is any AI-generated or manipulated image, audio or video content that resembles an existing person, object, place, entity or event and which would falsely appear to a person to be authentic or truthful.

Four things have to be true together: the content (i) resembles something (ii) real, that (iii) is a person, object, place, entity or event, and (iv) the result would falsely seem authentic or truthful to the viewer (i.e. it is sufficiently realistic that an ordinary viewer could take it to be genuine, regardless of whether any deception was intended). Importantly, the person, object, place, entity or event does not have to correspond to something that actually exists. It is sufficient that it could plausibly exist, or could plausibly have existed, in reality. This makes the concept of a deepfake extremely broad, extending well beyond digital replicas of real individuals to a wide range of realistic but entirely fictional content, which can itself be a source of deception, manipulation and disinformation.

The Guidelines give some useful examples that are directly relevant to the advertising industry and show when the rules apply and don't apply in practice:

Caught by the rule: deepfake	Not caught by the rule: not a deepfake
An AI-generated video featuring an AI depiction of a celebrity influencer in an advertising or promotional context; a synthetic avatar of a company CEO in a corporate video; an AI-generated product image that makes a product look better, different or more capable than it really is in a way that could mislead the audience about its appearance, characteristics or use.	A real product shown against an AI-generated background or surrounding environment, provided the ad isn't likely to mislead the audience about the product's actual representation, characteristics and use; AI-generated talking animals or clearly fantastical scenarios used in a way no reasonable viewer would take as real; minor AI-assisted background extension, colour correction, re-scaling or arrangement of products used for aesthetic purposes in packaging or advertisement photography.

The dividing line the Guidelines draw is whether the AI use affects the audience's perception of the truth or authenticity of what's shown.

It's worth noting that this deepfake test is specific to image, audio and video content; a separate and narrower labelling rule applies to AI-generated text. Text labelling under Article 50(4) only applies where the text informs the public on matters of public interest, such as politics, public health, consumer safety or environmental claims, not routine advertising or product description copy. In practice, this means most AI-written ad copy, headlines and search ads, including in formats like search and AdWords-style campaigns, should sit outside this particular obligation, unless the copy itself makes the kind of substantive public-interest claim described above, for example a specific health or sustainability claim. However, it is worth noting that there is a separate obligation on the provider of an AI system to mark AI-generated content (including text) as such, using methods that are detectable by machine-readable means (such as watermarking). For example, Anthropic has published [details of how it is now watermarking content produced using Claude](#).

In practice, if you've used AI to make the colours in your ad more vivid, sharpen an image, or clean up a background, you're likely fine. However, if you've used AI to create people, characters, products, or audiences that appear in the ad as if they were real, you are very likely to need to apply a label.

The lighter regime for evidently artistic, creative, satirical, fictional or analogous work

Where a deepfake forms part of an evidently (to the people exposed to it) artistic, creative, satirical, fictional or analogous work, the transparency obligation isn't completely switched off, but it is reduced. Disclosure is still required, but it is limited to the disclosure of the deepfake "in an appropriate manner that does not hamper the display or enjoyment of the work".

The Guidelines are clear that this is a narrow exception. If an ad is obviously just informative or commercial in nature i.e., its purpose is clearly to inform or sell, not to entertain or make an artistic point, it will not qualify for this lighter regime, even if it's creatively made.

The Guidelines give some useful examples that show when the lighter regime for evidently artistic, creative or satirical works is likely to apply in practice:

Qualifies for the lighter regime	Does not qualify for the lighter regime
----------------------------------	---

A movie or movie trailer featuring AI-generated or de-aged digital replicas of existing or deceased actors, shown in a cinema or on a streaming platform; AI-generated music in the style of an existing artist; an AI-manipulated image of a politician placed in a scene clearly meant to criticise, in a humorous way, a policy decision they took; AI-generated gaming imagery involving deepfake simulations of real people.	A photorealistic, AI-manipulated "teleshopping-style" ad depicting people using and endorsing a product, aimed at persuading viewers to buy it; an AI-generated image of celebrities implying they were involved in events that never happened, with no fictional, satirical or comparable purpose; an AI-manipulated video of a realistic synthetic influencer testing a sponsored product, focused purely on showing off its features; AI-generated video depicting realistic scenes of real historical atrocities shared on public social media.
---	---

In practice, the safe assumption for most commercial advertising will be to treat the full labelling obligation as the default. The lighter regime should only be relied on where the content is unambiguously and evidently artistic, fictional or satirical.

The AI icons – what they are and when to use them

The EU Commission has published a set of example icons, contained in the Code of Practice, that deployers can use to label AI-generated or manipulated deepfake content. They come in black, white, and 50%-transparent black/white variants, in SVG and PNG formats.

Their use is optional, with brands having the opportunity to design an equivalent label instead, provided it meets the same design and placement standards. There are three icon variants, each for a different scenario.

Below is a representation of the EU Commission's icons, and how they might apply in the context of advertising.

Icon	When to use it	Advertising example
Basic AI icon 	AI was involved in creating deepfake image, audio or video, or where a custom text label or interactive second layer (e.g. click for more information) is implemented.	Deepfake video with the text label "voices generated with" followed by the basic icon.
Fully AI-Generated icon 	The entire deepfake (image, audio or video) or text is fully generated by AI, with no human-created content or human editorial control beyond prompting.	A fully AI-generated advert scene, or an ad using fully AI-composed music or art.

Partially AI-Modified icon 	Content that started out as human-made (e.g. a photo, a video, or a piece of text) but has since been partially edited or altered using AI, to the point where it now counts as either a deepfake, or as AI-manipulated text on matters of public interest.	A real product photo where AI has swapped in a different setting or face, or furnished an empty room in a property or lifestyle ad.
--	---	---

How should the AI label be displayed?

For media and advertising businesses, the placement specifications translate into the below checklist based on the guidance in the Code of Practice:

- Show the icon or label at the very start of the video or image and, for longer or live content, at regular intervals and after every ad break (not just once at the outset).
- Embed it directly in the content itself so it survives reshares, screenshots, downloads and clipped fragments – rather than relying solely on a platform overlay that can be stripped out.
- Place it somewhere with no competing overlays (e.g., top corner of frame), keep it a clearly visible size, and make sure it stays legible against the background.
- Keep audible disclosures to a supporting role for visual deepfakes – never rely on sound alone, and always pair it with a visual disclosure (e.g. where audio is part of the user experience, or for accessibility purposes).
- When a screen is available, add a visual disclosure alongside an audible disclaimer.
- For audio-only ads (radio, podcasts, smart speakers), use a short, plain-language spoken disclaimer at the start, repeated after interruptions such as ad breaks.

The icons should be presented in a clearly visible size, and any accompanying label should use plain language and avoid jargon.

What steps should brands and agencies take now?

- Map which current and pipeline campaigns use AI-generated or manipulated imagery, audio or video, and screen each against the deepfake test.
- Separate genuinely cosmetic AI editing (which needs no AI label) from AI use that changes what the audience sees as real (which does need an AI label).
- Build the AI question into the creative brief from the outset, rather than raising it once creative is signed off: does this concept use AI, could it use AI, and if so, would a label change the message or the way it needs to be shot or produced.
- If you're an agency, treat any proposed use of AI as an automatic trigger to consider the labelling requirement, in the same way a brief would trigger a rights or clearance check.
- If you're a brand working with an agency, make "has AI been used, and if so how" a standard question on every campaign update or creative review, not just a one-off check at the start.

- Build the EU icons, or an equivalent label of your own design, into the advert at the placement points set out above.

It is also worth bearing in mind that AI labelling may not be everything you need to do to protect a brand from liability more broadly. Adding an AI label does not clear a deepfake of separate legal duties, which still apply in full. For example, if the deepfake uses someone else's trade mark, copyright, or other protected material, intellectual property law will apply, and an AI label has no effect on whether that material is protected. Separately, where a deepfake shows a real person, rights in their image or voice may still apply too (which may lead to data protection concerns, as well as arguments of passing off). Labelling does not make otherwise unlawful content lawful.

Why this is important

Penalties for non-compliance with Article 50 can reach up to €15 million or 3% of global annual turnover for both providers and deployers, whichever is higher, so compliance should be a serious priority for advertisers and their agencies.

There is also a reputational dimension. The rules are intended to prevent consumers being deceived by some kinds of AI-generated content and to protect public trust. Advertising depends on that same trust, so brands that are perceived to be misleading their customers with undisclosed AI deepfakes risk more than a fine – they risk jeopardising the trust their advertising relies on.

What is the position in the UK?

Many UK businesses will be in scope of the EU AI Act's reach, for the reasons set out above. In the UK, there is no current regulation relating to labelling of AI-generated content. Of course, general rules around misleading advertising still apply under the CAP Codes, as do existing laws relating to IP and data protection. However, following its consultation on the interplay between copyright and AI, the [Government has indicated](#) that it will be issuing a standalone consultation on labelling of digital replicas (deepfakes). This was expected to be published in summer 2026, so developments should be closely monitored.

Authors: [Ana-Maria Curavale, Associate](#), and [Stuart Lester, Partner](#), Mishcon de Reya LLP