

NO TODO CONTENIDO GENERADO CON IA DEBE ETIQUETARSE

JON SANMILLÁN FUICA

Guía sobre cómo aplicar el artículo 50.4 del AI Act a las ultrafalsificaciones y a los textos de interés público

AI GOVERNANCE · EDICIÓN DE LA GUÍA: 9 DE SEPTIEMBRE DE 2026

RESUMEN. El artículo 50.4 del Reglamento de Inteligencia Artificial obliga a determinados responsables del despliegue a revelar el origen artificial de dos clases de contenidos: las imágenes, audios o vídeos que constituyan una ultrafalsificación y ciertos textos publicados para informar al público sobre asuntos de interés público. El objetivo es que las personas cuenten con información sobre la procedencia del contenido y así puedan ajustar sus decisiones y confianza. Esta obligación resulta incómoda para los usuarios puesto que supone un esfuerzo en un ámbito en el que prima la inmediatez y, lo más importante, mitiga el impacto que genera el contenido y su valoración por la sociedad. Dicho esto, la obligación no alcanza a todo contenido en el que haya intervenido una IA. Para determinar cuándo se aplica hay que examinar quién utiliza el sistema, qué contenido se ha generado o manipulado, cómo puede percibirlo su audiencia y si concurre alguno de los regímenes especiales o excepciones previstas. Este artículo convierte los criterios de las Directrices de la Comisión y el reciente Código de buenas prácticas para las obligaciones de transparencia del artículo 50 en un método de análisis y cumplimiento aplicable a empresas, medios, agencias, administraciones y profesionales.

CONTENIDO

1. Una norma con dos obligaciones diferentes	3
2. Antes de analizar el contenido: ¿quién es responsable del despliegue?	5
3. Primera obligación de transparencia: cuándo una imagen, un audio o un vídeo es una ultrafalsificación y cuándo debe ser etiquetada	6
4. Segundo deber de transparencia: qué textos sobre asuntos de interés público deben etiquetarse.....	10
5. Cómo debe presentarse la información	14
6. Matriz de supuestos y método de cumplimiento para organizaciones y profesionales	17
7. Aplicación temporal, supervisión y sanciones	23
8. Límites del artículo 50.4	24
9. Conclusiones	25
Referencias	26

Tres decisiones antes de actuar

1. Identificar quién actúa como responsable del despliegue y si el contenido se utiliza con fines profesionales.
2. Clasificar el resultado: ultrafalsificación, texto de interés público o supuesto situado fuera de esas dos ramas.
3. Elegir una divulgación adecuada al contexto y conservar evidencia proporcionada de la revisión, la decisión y el funcionamiento.

Límites. La guía ofrece orientación general y no sustituye el análisis del caso concreto. El etiquetado no acredita la veracidad, la autoría o la licitud del contenido ni agota otras obligaciones aplicables. La saturación de etiquetas puede llevar a ignorarlas y generar confianza inversa: interpretar la ausencia de etiqueta como garantía de autenticidad. La excepción editorial también plantea un límite probatorio: una política general no siempre acredita la revisión de una publicación concreta, pero exigir un expediente exhaustivo para cada texto puede imponer cargas desproporcionadas e interferir en los procesos editoriales.

Un vídeo muestra a una persona semejante a un político pronunciando un discurso que nunca dio. Un informe de sostenibilidad elaborado con ayuda de inteligencia artificial ha sido revisado por profesionales de cumplimiento antes de publicarse. En un anuncio, unos ratones generados mediante IA discuten, en lenguaje humano, sobre cuál es el mejor queso.

En los tres contenidos ha intervenido una IA. Sin embargo, la respuesta jurídica no es la misma.

Las [Directrices de la Comisión sobre las obligaciones de transparencia del artículo 50](#) consideran que el primer vídeo puede constituir una ultrafalsificación y debe identificarse como artificial; que el informe puede quedar exento de etiquetado si la revisión ha sido sustantiva y existe responsabilidad editorial; y que el anuncio de los ratones no constituye una ultrafalsificación porque representa una escena manifiestamente irreal sin capacidad para inducir a error sobre su autenticidad.

Este contraste resume la idea central del artículo 50.4 del Reglamento de Inteligencia Artificial: **no todo contenido generado o manipulado mediante IA debe etiquetarse de forma perceptible para el público en virtud de este apartado**. Tampoco basta con preguntar si la IA ha intervenido. Hay que identificar el actor, la modalidad del contenido, su función, el contexto de publicación y las expectativas de la audiencia.

¿Por qué son importantes las obligaciones de transparencia del artículo 50.4? El artículo 50.4, junto con el resto del contenido del artículo 50, es aplicable desde el 2 de agosto de 2026 (excepto el periodo de transición para las obligaciones del artículo 50.2). No obstante, este apartado resulta especialmente relevante, dado que afecta directamente a los responsables del despliegue y a dos de los casos de uso más comunes de los sistemas de IA en la actualidad: la generación o manipulación de textos y ultrafalsificaciones.

Su alcance no se entiende únicamente leyendo el precepto. Buena parte de las dudas se encuentran hoy matizadas en las Directrices publicadas por la Comisión el 20 de julio de 2026. Estas Directrices no son vinculantes y solo el Tribunal de Justicia de la Unión Europea puede ofrecer una interpretación definitiva del Reglamento, pero constituyen la

principal referencia institucional para autoridades, proveedores y responsables del despliegue mientras se desarrolla la práctica supervisora.

1. UNA NORMA CON DOS OBLIGACIONES DIFERENTES

El [artículo 50.4](#) contiene dos obligaciones dirigidas a los **responsables del despliegue** de sistemas de IA:

1. Quien utilice un sistema de IA para generar o manipular imágenes, audio o vídeo que constituyan una **ultrafalsificación** debe revelar que el contenido ha sido generado o manipulado artificialmente.
2. Quien utilice un sistema de IA para generar o manipular un **texto publicado con la finalidad de informar al público sobre asuntos de interés público** debe revelar su origen artificial, salvo que se cumpla la excepción de revisión humana o control editorial y responsabilidad editorial.

Aunque ambas obligaciones persiguen que las personas reconozcan el origen artificial del contenido, no funcionan de la misma manera.

Tabla 1. Comparación entre ultrafalsificaciones y textos de interés público.

Elemento	Ultrafalsificaciones	Textos de interés público
Modalidades	Imagen, audio o vídeo	Texto
Presupuesto principal	El contenido reúne la definición de deepfake (ultrafalsificación) del artículo 3.60	El texto se publica para informar al público sobre un asunto de interés público
Regla	Debe revelarse que el contenido ha sido generado o manipulado artificialmente	Debe revelarse que el texto ha sido generado o manipulado artificialmente
Régimen especial	Divulgación atenuada para obras evidentemente artísticas, creativas, satíricas, ficticias o análogas	No existe un régimen artístico equivalente dentro de esta rama
Excepción más relevante	Uso autorizado por ley para determinadas finalidades de aplicación de la ley	Revisión humana o control editorial, junto con responsabilidad editorial; además de la excepción para determinadas finalidades de aplicación de la ley

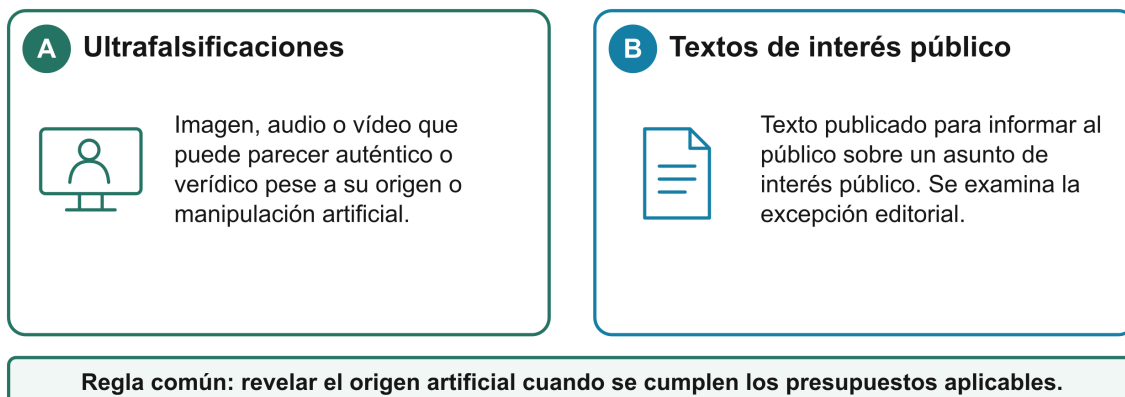


FIGURA 1. Dos clases de contenidos sujetos a la obligación de divulgación del artículo 50.4. Los ejemplos son esquemáticos; las condiciones y los regímenes especiales se desarrollan en el texto.

Esta separación es imprescindible. La revisión humana puede excluir del etiquetado a determinados textos, pero no constituye una excepción equivalente para una ultrafalsificación audiovisual. Del mismo modo, el hecho de que una imagen haya sido creada íntegramente con IA no permite concluir por sí solo que sea una ultrafalsificación.

También debe distinguirse el **etiquetado perceptible** del artículo 50.4 del **marcado técnico** del artículo 50.2. El proveedor del sistema debe incorporar, cuando proceda, soluciones legibles por máquina que permitan detectar el origen artificial del resultado, aunque los usuarios no lo noten. Sin embargo, el responsable del despliegue debe informar de manera comprensible a la persona expuesta a una ultrafalsificación o a uno de los textos cubiertos. Las obligaciones pueden acumularse y corresponder a sujetos distintos. Un metadato o una marca técnica que la audiencia no puede percibir sin herramientas específicas no sustituye la divulgación del artículo 50.4.

1.1. ¿Qué pretende evitar el artículo 50.4?

Los considerandos 132 a 136 del AI Act atribuyen a las obligaciones del artículo 50 el objetivo de que las personas puedan **reconocer cuándo una interacción o un contenido se encuentra mediado por IA para ajustar su confianza y tomar decisiones informadas**.

El considerando 133 se centra en la dificultad creciente para distinguir el contenido sintético del contenido auténtico y justifica la utilización de técnicas de marcado y detección. El considerando 134 justifica la transparencia sobre las ultrafalsificaciones y determinados textos de interés público, pero introduce mecanismos de proporcionalidad para no restringir indebidamente la libertad de expresión ni las libertades artística y científica. Por su parte, el considerando 136 identifica como posibles consecuencias de la generación de este tipo de contenido los efectos negativos reales o previsibles en los procesos democráticos, el discurso cívico y los procesos electorales, incluso a través de la desinformación.

Así pues, el artículo 50 y, en particular, su apartado 4 nacen para hacer frente a posibles peligros como **la suplantación, el engaño, la desinformación, la manipulación y el fraude**. Todos estos peligros pueden tener grandes impactos en la sociedad, como ejemplifica el considerando 136. Para ello, se pretende dotar a las personas de información acerca de la procedencia real del contenido.

No obstante, al tratar de afrontar estos riesgos, el precepto genera una serie de tensiones con otros derechos, como la libertad de expresión e información, ya que el etiquetado condiciona la forma en la que una persona comunica su contenido; o la propiedad intelectual y el derecho moral del autor, dado que determinadas etiquetas mitigan la valoración popular del contenido, pudiéndose llegar al extremo de entender que este carece de autor.

A su vez, es preciso identificar que el artículo 50.4 tiene un objeto limitado. **La etiqueta informa del origen artificial o de la manipulación del contenido; no certifica su veracidad, fiabilidad, licitud ni calidad**. Un vídeo puede estar etiquetado correctamente y seguir siendo difamatorio, vulnerar derechos de imagen o constituir publicidad engañosa. Esa distinción es importante para evitar una expectativa equivocada: el artículo 50.4 no crea un sistema general para determinar qué contenidos son verdaderos. Crea una señal que permite al destinatario valorar el contenido sabiendo que ha sido generado o manipulado artificialmente. Trataremos esta idea junto con otras, como el riesgo de fatiga y de confianza inversa, la exclusión de determinados contenidos o la dificultad probatoria de la excepción editorial, a lo largo de este artículo.

2. ANTES DE ANALIZAR EL CONTENIDO: ¿QUIÉN ES RESPONSABLE DEL DESPLIEGUE?

El primer paso no consiste en examinar la imagen o el texto. Consiste en determinar quién utiliza el sistema de IA bajo su autoridad.

Según las Directrices, **tener autoridad** sobre el sistema significa **asumir la responsabilidad sobre la decisión de utilizarlo y sobre la forma y finalidad de ese uso, incluidos sus resultados**. No es necesario disponer de control técnico sobre el funcionamiento del modelo. Una empresa puede ser responsable del despliegue aunque utilice una herramienta proporcionada por un tercero, siempre que decida para qué y cómo emplea sus resultados.

2.1. Empresas, trabajadores, agencias y profesionales externos

Cuando una persona jurídica utiliza el sistema bajo su autoridad, **sus empleados no se convierten en responsables del despliegue independientes** por generar una imagen, redactar un texto o editar un vídeo siguiendo sus instrucciones. La organización conserva ese rol.

Lo mismo puede suceder cuando intervienen contratistas o profesionales externos que operan por cuenta de la empresa, bajo su responsabilidad y control. Sin embargo, las Directrices ofrecen un matiz relevante: una empresa que se limita a encargar una campaña a una agencia, sin decidir ni controlar si la agencia utiliza IA o cómo lo hace, no se considera responsable del despliegue por ese solo hecho. En ese supuesto, habrá que analizar la posición de la propia agencia.

La pregunta práctica no es, por tanto, quién pulsó el botón de generar. Es quién tomó la decisión sobre la finalidad y la forma de utilizar el sistema.

2.2. Plataformas, medios y otros distribuidores

Una plataforma, un servicio de alojamiento o un medio que se limita a transmitir o difundir contenido generado por un tercero no es responsable del despliegue si no tuvo autoridad sobre el sistema utilizado para crearlo o manipularlo. Las Directrices les animan a preservar las etiquetas y los marcados existentes, pero esa recomendación no transforma automáticamente al distribuidor en responsable del despliegue.

La situación cambia si la propia plataforma o el medio utiliza un sistema bajo su autoridad para generar contenidos. **También pueden resultar aplicables obligaciones distintas, como las medidas de mitigación de riesgos y las exenciones de responsabilidad relativas al fenómeno del "Safe Harbour" del Reglamento de Servicios Digitales (Digital Services Act o DSA)** para plataformas y buscadores de muy gran tamaño.

2.3. Uso personal y actividad profesional

El AI Act excluye las obligaciones de los responsables del despliegue cuando una persona física utiliza un sistema en una actividad **puramente personal y no profesional**. Las Directrices incluyen como ejemplo a una persona que genera y comparte una ultrafalsificación en redes sociales a título estrictamente personal.

La exclusión desaparece cuando existe una actividad empresarial, comercial, profesional, laboral o autónoma, o cuando la persona obtiene regularmente un beneficio económico. Por ello, un creador de contenido que monetiza de forma habitual sus publicaciones no debería asumir que todo lo que hace en sus perfiles constituye uso personal.

2.4. Responsables situados fuera de la Unión

El artículo 50.4 **puede alcanzar a responsables situados en terceros países cuando prevean que los resultados del sistema se utilizarán en la Unión**. Las Directrices mencionan como ejemplo la publicación de una ultrafalsificación en una página globalmente accesible cuando su difusión en la Unión ha sido dirigida o autorizada. No se extiende, en cambio, a una llegada a público europeo que sea imprevisible y se encuentre fuera del control del responsable.

3. PRIMERA OBLIGACIÓN DE TRANSPARENCIA: CUÁNDO UNA IMAGEN, UN AUDIO O UN VÍDEO ES UNA ULTRAFALSIFICACIÓN Y CUÁNDO DEBE SER ETIQUETADA

El artículo 3.60 define la ultrafalsificación como un contenido de imagen, audio o vídeo generado o manipulado por IA que se asemeja a personas, objetos, lugares, entidades o sucesos reales y que puede inducir a una persona a pensar erróneamente que es auténtico o verídico.

Las Directrices convierten esta definición en un análisis de varios pasos.



FIGURA 2. Desglose didáctico de la definición de ultrafalsificación y papel del contexto. El régimen de las obras evidentemente artísticas adapta la divulgación; no elimina el deber de informar.

3.1. Presupuesto previo: intervención de un sistema de IA y con fines profesionales

El contenido debe haber sido generado o manipulado mediante un sistema de IA utilizado con fines profesionales. La generación parte de material sintético producido por el sistema; la manipulación altera contenido ya existente, sea humano o sintético.

No es necesario que todo el contenido haya sido creado por IA. Una modificación parcial puede quedar incluida si afecta a elementos relevantes para la percepción del destinatario.

"Caso Práctico"

Una persona comparte en redes sociales personales un deepfake sobre un político diciendo cosas que no ha dicho. Paradójicamente, este hecho, aunque cumpla con el supuesto de hecho del artículo 50.4, puede no verse afectado por las obligaciones de transparencia descritas en este artículo. El Reglamento no persigue la actividad estrictamente personal de los sujetos; esto implica que se puedan dar situaciones en las que un particular por diversión o activismo cree este tipo de deepfakes que se vuelvan virales, teniendo consecuencias negativas como la desinformación; y que no esté obligado a etiquetar el contenido.

Dicho esto, habrá que atender a la naturaleza de la cuenta y a la actividad de la persona que publica el contenido. Si la cuenta se monetiza regularmente o concurren una actividad de influencer, un trabajo autónomo, una campaña publicitaria, un encargo remunerado o una actuación por cuenta de un partido, medio, empresa o agencia, estaremos ante supuestos de aplicación del Reglamento de IA y de su artículo 50.4.

3.2. Primer criterio: semejanza apreciable

Debe existir un **alto grado de similitud entre el contenido y el sujeto representado**. No se exige identidad absoluta. La comparación debe atender a rasgos característicos o distintivos que permitan reconocer a la persona, objeto, lugar, entidad o acontecimiento simulado.

Por eso puede constituir una ultrafalsificación tanto la réplica digital de una persona concreta como una voz que reproduzca de forma reconocible a un presentador o un vídeo que muestre a dos futbolistas reales en una escena que nunca ocurrió.

3.3. Segundo criterio: un sujeto existente o plausible

Las Directrices interpretan "existente" de forma más amplia de lo que podría sugerir una lectura literal. No solo incluye aquello que existe actualmente. También comprende personas, objetos, lugares, entidades o sucesos que **pueden existir plausiblemente o podrían haber existido en la realidad**.

Esto tiene una consecuencia relevante: una persona sintética que no copia a ningún individuo concreto puede constituir una ultrafalsificación si parece un ser humano real y puede inducir a la audiencia a creer que existe. Las Directrices incluyen dentro del concepto de persona a las réplicas digitales de individuos reales, pero también a avatares o personajes humanos sintéticos realistas.

Quedan fuera las representaciones manifiestamente imposibles y carentes de potencial para engañar, como un dragón, una persona volando sin ayuda mecánica o un elefante conduciendo un coche.

3.4. Tercer criterio: una de las categorías previstas

La simulación debe referirse a una persona, objeto, lugar, entidad o suceso:

- **Personas:** seres humanos realistas, réplicas digitales, avatares humanos plausibles, imagen, voz, comportamiento o actuación.
- **Objetos:** bienes o elementos materiales realistas, incluidos productos, edificios, obras de arte o maquinaria.
- **Lugares:** localizaciones realistas.
- **Entidades:** seres vivos no humanos plausibles, como animales.
- **Sucesos:** escenas o situaciones realistas, también cuando recrean acontecimientos históricos o la prestación de un servicio.

Esta amplitud explica por qué el precepto no se limita a la suplantación de celebridades. Puede afectar a la imagen artificial de un producto, a la recreación de un acontecimiento o a un avatar corporativo que aparente ser una persona real.

3.5. Cuarto criterio: falsa apariencia de autenticidad o veracidad

Este es el criterio decisivo y el lugar en el que las Directrices reducen gran parte de la ambigüedad del Reglamento.

El contenido debe **tener capacidad para engañar o inducir a error** sobre alguna de las siguientes características:

- **Su autenticidad:** si el contenido es realmente lo que aparenta ser respecto de su origen o proceso de creación. Por ejemplo, si intervino una persona real, si pronunció esas palabras o si el producto tiene realmente ese aspecto.
- **Su veracidad:** si la representación es correcta desde el punto de vista factual.

No hace falta que el responsable quiera engañar. La evaluación es objetiva. Un contenido puede ser una ultrafalsificación aunque su creador no tuviera intención fraudulenta si, atendiendo al contexto, puede parecer auténtico o verdadero.

Para realizar la evaluación deben considerarse conjuntamente:

1. El **nivel de semejanza**.
2. El **mensaje sustantivo** que transmite el contenido.
3. El **contexto y el entorno de difusión** previstos y previsibles.
4. La **composición de la audiencia razonablemente previsible** y sus expectativas.

Las Directrices no utilizan aquí la figura de una persona media razonablemente informada que aparece en otros apartados del artículo 50. Exigen considerar la diversidad real de la audiencia. Si es previsible que el contenido llegue a niños, personas mayores o colectivos con menor alfabetización digital, basta con que esa parte de la audiencia pueda resultar engañada para que el contenido llegue a calificarse como ultrafalsificación.

El responsable no tiene que anticipar cualquier difusión remota e imprevisible por terceros. Si el contenido se comparte exclusivamente en una zona para suscriptores o en un boletín corporativo, no debe presumirse automáticamente una exposición pública general. Sí debe considerarse, en cambio, la audiencia que resulte razonablemente previsible a partir del canal y del uso elegidos.

Por ejemplo, un vídeo hiperrealista de un político puede no constituir una ultrafalsificación cuando se muestra exclusivamente en un entorno formativo cerrado a especialistas que saben que están examinando simulaciones sintéticas. No se trata de una excepción para audiencias alfabetizadas: falta el requisito de que el contenido pueda parecerles falsamente auténtico o verídico.

3.6. El contexto decide: supuestos que las Directrices incluyen y excluyen

La principal aportación práctica de las Directrices es que el realismo visual o sonoro no basta por sí solo. La misma técnica puede producir un contenido sujeto a etiquetado en un contexto y un contenido fuera de la definición en otro.

3.6.1. *Contenidos que sí pueden constituir ultrafalsificaciones*

La Comisión menciona, entre otros, los siguientes ejemplos:

- Una imagen manipulada que muestra a dos futbolistas profesionales reales frente a un estadio.
- Un audio que clona las voces de los presentadores habituales de un podcast periodístico y de una persona invitada para hablar sobre noticias.
- Un vídeo de una persona semejante a un político pronunciando un discurso ante una audiencia.
- Una celebridad o una persona influyente sintética utilizada en publicidad.
- Un avatar humano realista del consejero delegado de una empresa felicitando a los trabajadores por los resultados anuales.
- Una imagen artificial de un producto que altera la percepción sobre su aspecto, características, calidad o utilización real.

El último ejemplo es especialmente importante. El concepto no queda reservado a la desinformación política o a la imitación de voces. Una presentación artificial de un producto puede ser una ultrafalsificación cuando altera de forma relevante lo que el consumidor cree que está viendo.

3.6.2. *Contenidos que, según las Directrices, no son ultrafalsificaciones*

Entre los ejemplos excluidos aparecen:

- Una esfinge volando sobre la Torre Eiffel.
- Unos ratones que hablan sobre queso en un anuncio.
- La voz artificial de un personaje ficticio cuando no existe engaño sobre la identidad del narrador.
- La normalización del volumen, reducción de ruido o compresión de un programa de radio sin alterar las palabras ni la forma de hablar.
- Una representación en forma de dibujo animado de un acontecimiento histórico.
- Entornos ficticios generados para videojuegos.
- Actores reales, no manipulados, que actúan sobre un fondo artificial claramente ficticio.
- Un producto real mostrado sobre un fondo generado mediante IA, siempre que el anuncio no induzca a error sobre el producto, sus características o su uso.

No deben convertirse estos ejemplos en exclusiones automáticas. El fondo artificial de un anuncio puede ser irrelevante cuando cumple una función estética, pero una modificación sustancial del fondo de una fotografía periodística puede cambiar la interpretación del acontecimiento y afectar a la percepción de autenticidad.

3.6.3. *Ajustes técnicos y modificaciones menores*

Las Directrices consideran que **determinadas modificaciones de escasa relevancia material pueden no convertir el contenido en una ultrafalsificación**. Citan la eliminación

de transeúntes del fondo, los ajustes de iluminación, la corrección de color, la reducción de ruido, la compresión, ciertos retoques cosméticos o las ampliaciones estéticas de un fondo.

El criterio no es el nombre comercial de la función utilizada, sino su impacto. Una corrección de color aplicada a la fotografía de un producto puede ser neutra; esa misma intervención, utilizada para cambiar sustancialmente una imagen periodística, puede alterar la forma en que el público entiende lo sucedido.

La pregunta correcta es: **¿la intervención cambia de forma relevante lo que la audiencia puede creer sobre el origen, la realidad o el significado del contenido?**

3.7. Obras artísticas, creativas, satíricas o ficticias: obligación atenuada, no exención

Cuando una ultrafalsificación forma parte de una obra o programa **evidentemente artístico, creativo, satírico, ficticio o análogo**, el artículo 50.4 permite revelar su existencia de una manera apropiada que no perjudique la exhibición o el disfrute de la obra.

Este régimen suele describirse erróneamente como una excepción. No lo es. El contenido sigue sujeto a transparencia; lo que se flexibiliza es la forma de cumplirla.

Las Directrices exigen interpretar de manera estricta la palabra "evidentemente". La naturaleza artística, creativa, satírica o ficticia debe resultar reconocible para las personas expuestas. Para valorarlo pueden considerarse:

- El estilo o formato del contenido, como la exageración o la ironía propias de la sátira.
- El lugar y el medio en el que se presenta.
- Las expectativas de la audiencia, por ejemplo, en una película, un videojuego o una experiencia de realidad virtual.

Si la naturaleza del contenido resulta ambigua, no procede aplicar el régimen atenuado. Tampoco basta con invocar creatividad para cubrir una finalidad exclusivamente informativa o comercial. Las Directrices señalan que algunos anuncios o documentales podrían considerarse creativos o ficticios en contextos concretos, pero no de forma automática. Cuando un contenido combine una función informativa con otra creativa, debe prevalecer la función informativa y aplicarse el etiquetado ordinario.

Entre los ejemplos que pueden acogerse al régimen atenuado aparecen una película que rejuvenece digitalmente a un actor, la réplica de un actor fallecido, una imagen satírica de un político o un videojuego que representa a personas reales. En cambio, un vídeo de televenta con consumidores sintéticos mostrando el funcionamiento de un producto, una celebridad asociada a hechos inexistentes sin propósito satírico o un *influencer* sintético que prueba un producto patrocinado se someten al régimen ordinario.

Este régimen permite que, según el Código de buenas prácticas, **el etiquetado pueda adaptarse al contexto**: una indicación junto al contenido, una nota descriptiva, una interfaz visible o, en determinadas obras, los elementos de presentación o créditos. En esencia, se permite establecer este tipo de etiquetas de manera exterior al contenido, pero de forma que sigan vinculadas a él. Por ejemplo, pueden situarse fuera, pero junto al marco de la imagen o del vídeo; en las notas o descripción de la obra; en la entrada física; o en el punto de venta.

Recurso ejecutivo descargable. [Descargar «Ultrafalsificaciones: cuándo revelar el uso de IA»](#). PDF · DIN A4 vertical · 1 página · 96 kB · Edición del 9 de septiembre de 2026.

4. SEGUNDO DEBER DE TRANSPARENCIA: QUÉ TEXTOS SOBRE ASUNTOS DE INTERÉS PÚBLICO DEBEN ETIQUETARSE

El artículo 50.4 no obliga a identificar cualquier texto redactado con ayuda de IA. Para entrar en su segunda rama deben concurrir tres elementos: **que el texto esté publicado, que tenga la finalidad de informar al público y que trate sobre un asunto de interés público.**



FIGURA 3. Los requisitos de la rama textual y las dos condiciones acumulativas de la excepción editorial. La revisión humana y el control editorial son alternativas dentro de la primera condición; la responsabilidad editorial se exige además.

4.1. El texto debe estar publicado

Según las Directrices, un texto está publicado cuando pueda **acceder a él un número indeterminado y suficientemente amplio de personas no relacionadas, de forma simultánea o sucesiva**. No importa que el acceso sea gratuito o de pago: un artículo sujeto a suscripción puede estar publicado.

Quedan fuera:

- La correspondencia privada, también cuando sea profesional.
- Un informe dirigido exclusivamente a un cliente.
- Las comunicaciones internas de una organización.
- Un grupo privado y reducido de mensajería.
- Una respuesta o resumen generado por un chatbot que solo puede ver la persona que lo solicitó.

4.2. Debe tener la finalidad de informar al público

Las directrices puntualizan que el **texto debe pretender comunicar "conocimientos, opiniones o hechos"**. Las expresiones breves que no transmiten materialmente este tipo de contenido no se consideran, por sí solas, información al público.

Este requisito permite diferenciar una publicación informativa de una etiqueta de interfaz, una frase ornamental o una comunicación puramente funcional. Lo relevante vuelve a ser la finalidad y el contenido real, no el formato utilizado.

4.3. Debe referirse a un asunto de interés público

Las Directrices entienden por asuntos de interés público aquellos **asuntos relevantes para la sociedad en general, a escala local, nacional, europea o internacional, y que merecen debate o escrutinio público**. Este es el requisito más ambiguo de los que se exigen para este tipo de contenido. Las directrices incluyen la siguiente lista abierta de posibles supuestos de interés público:

- Política y procesos democráticos.
- Administración y servicios públicos.
- Administración de justicia y aplicación de la ley.
- Derechos fundamentales.
- Seguridad pública.
- Salud pública.
- Protección medioambiental.
- Seguridad de los consumidores.
- Desarrollos económicos, financieros, políticos, científicos o culturales relevantes para el debate público.

El interés público es contextual y puede evolucionar. No equivale a aquello que genera curiosidad o tráfico. Una noticia sobre una decisión municipal puede ser de interés público local aunque no tenga alcance nacional. Una descripción comercial ordinaria puede quedar fuera, pero una afirmación sobre los efectos de un producto en la salud, su seguridad o su sostenibilidad puede introducir el texto en una materia de interés público.

4.4. Ejemplos oficiales incluidos y excluidos

Las Directrices consideran comprendidos, salvo que resulte aplicable la excepción editorial:

- Un resumen generado mediante IA de un artículo periodístico sobre una decisión municipal.
- Partes manipuladas mediante IA de un artículo que compara el efecto de distintas dietas sobre una enfermedad.
- Un informe corporativo de una sociedad cotizada que contiene información para inversores.
- Un mensaje generado mediante IA por un instituto meteorológico para advertir sobre una tormenta y comunicar medidas de precaución.

Consideran fuera del supuesto:

- Una novela de fantasía.
- Un anuncio o una descripción ordinaria de producto que no contenga afirmaciones sobre salud, seguridad de consumidores o sostenibilidad.

- Un resumen de noticias generado por un chatbot únicamente para quien lo solicita.
- Un texto elaborado por un consultor para asesorar privadamente a un cliente sobre cumplimiento normativo.

La comparación muestra que el precepto no se limita a medios de comunicación. Puede alcanzar a empresas cotizadas, organismos públicos, centros de investigación, asociaciones, consultoras o profesionales que publiquen contenidos informativos sobre materias relevantes para el debate público.

4.5. La excepción editorial: revisar no significa simplemente leer

Un texto comprendido en el artículo 50.4 no necesita etiquetarse cuando se cumplen **dos condiciones acumulativas**:

1. El contenido ha sido **sometido a revisión humana o control editorial**.
2. Una **persona física o jurídica asume la responsabilidad editorial** de la publicación.

No basta con una de ellas. Tener una política editorial sin aplicar una revisión sustantiva no es suficiente. Tampoco lo es revisar el contenido si nadie asume la responsabilidad última sobre su publicación.

4.5.1. *Qué es una revisión humana sustantiva*

La revisión humana consiste en un **examen deliberado de la sustancia del texto realizado por una o varias personas con conocimientos relevantes y criterio profesional sobre la materia**. Las Directrices establecen que la comprobación de la exactitud de los hechos debe formar parte, como mínimo, de esa revisión.

El control editorial se refiere al **poder ejercido en la práctica por una entidad responsable para aprobar, modificar o rechazar el contenido por razones sustantivas**, incluidas la comprobación de hechos y la fiabilidad de las fuentes.

No cumplen este estándar:

- La mera corrección ortográfica o gramatical.
- La existencia formal de una política editorial sin pruebas de aplicación práctica.
- Una revisión realizada exclusivamente por otro sistema de IA.
- Una aprobación rápida sin intervención sustantiva.
- Un procedimiento automático presentado como si fuera supervisión humana.

Además, si después de la revisión se utiliza nuevamente la IA para modificar, ampliar o reformular sustancialmente el texto, la excepción deja de resultar aplicable. La revisión debe recaer sobre la versión que finalmente se publica.

4.5.2. *Qué implica asumir responsabilidad editorial*

La persona física o jurídica debe asumir la **responsabilidad legal última sobre la publicación y sobre el proceso de revisión o control editorial**. Puede tratarse de un editor, un consejo editorial, una empresa editora, un organismo público o el profesional que publica bajo su propia responsabilidad.

Las Directrices recomiendan que **la identidad y los datos de contacto de la persona, entidad o función responsable sean fácilmente localizables**, por ejemplo, en la información legal de la web o en los datos editoriales de la publicación.

Los medios pueden apoyarse en sus procedimientos y estándares profesionales existentes cuando estos satisfagan los requisitos. Pero la excepción no está reservada a los

medios. Un centro de investigación, una empresa o una administración también pueden aplicarla si organizan una revisión competente y asumen realmente la responsabilidad editorial.

4.5.3. *¿Puede una misma persona revisar y asumir la responsabilidad?*

Las Directrices no exigen expresamente que el revisor y el responsable editorial sean **personas distintas**. Por tanto, cabe sostener que un profesional independiente puede reunir ambos papeles si posee conocimientos suficientes, revisa de manera deliberada la versión final, comprueba los hechos y asume públicamente la responsabilidad de publicación.

Esta es una interpretación razonada, no una exención automática. La mera circunstancia de que una persona publique un texto en su propio nombre no demuestra que haya existido revisión. De hecho, las Directrices consideran insuficiente un libro autopublicado sobre cambio climático generado mediante IA que no haya sido revisado por una persona competente.

La conclusión práctica es que un autor individual no necesita crear una redacción ficticia, pero sí **debe poder acreditar un proceso real**: contraste de fuentes, corrección sustantiva, decisión consciente sobre la versión final y responsabilidad identificable.

Por ejemplo, yo no tendría la necesidad de etiquetar estos artículos que realizo con el apoyo de la Inteligencia Artificial conforme a lo dispuesto en el artículo 50.4. Si bien es cierto que la IA me ayuda con la redacción de determinadas partes de estos artículos, yo posteriormente realizo una revisión sustantiva, llegando a modificar ampliamente el contenido que me sugiere la IA, adecuándolo a mi estilo y verificando íntegramente su veracidad. Además, son temas en los que poseo conocimientos relevantes que permiten una revisión sustantiva dado que cuento con certificación como Delegado de Protección de Datos, soy AI Compliance Professional, abogado colegiado no ejerciente y jurista especializado en privacidad y derecho digital.

4.5.4. *Ejemplos en los que puede aplicarse la excepción*

La Comisión menciona:

- Un artículo de prensa manipulado mediante IA o su resumen generado con IA, sometido al control del editor responsable.
- Un blog académico revisado internamente por pares y publicado bajo la responsabilidad del centro de investigación.
- Una alerta de seguridad pública generada con IA y aprobada por un funcionario competente bajo la responsabilidad del organismo de protección civil.
- Un informe de sostenibilidad revisado por profesionales de cumplimiento de una empresa cotizada.
- La traducción asistida por IA de un artículo humano cuando la traducción final ha sido revisada.

Por el contrario, deben etiquetarse, si concurren los demás requisitos, una web que publica automáticamente artículos sobre política de la Unión, un texto revisado por otra IA y sometido después a una mera corrección gramatical humana o una publicación autopublicada sin revisión competente.

5. CÓMO DEBE PRESENTARSE LA INFORMACIÓN

El artículo 50.5 completa las dos obligaciones del apartado 4: la información debe proporcionarse de forma **clara y distinguible**, como máximo en el momento de la **primera exposición de cada persona**, y respetar los **requisitos de accesibilidad aplicables**. El [Código de buenas prácticas sobre transparencia de los contenidos generados por IA](#) convierte ese estándar abierto en especificaciones de diseño, ubicación y organización.

Conviene diferenciar su fuerza jurídica. La adhesión al Código es voluntaria y no constituye una prueba concluyente de cumplimiento. Sin embargo, quienes se adhieren a su sección 2, sobre las obligaciones del artículo 50.4, se comprometen a aplicar las medidas expresadas como **"will"**; las formuladas como **"encouraged"** o **"may"** son recomendaciones o posibilidades voluntarias. Un responsable no adherido puede utilizar otra solución, pero deberá demostrar que satisface los artículos 50.4 y 50.5 por medios igualmente adecuados.

5.1. Un identificador reconocible y una explicación precisa

Cuando sea posible informar visualmente, el Código exige a sus signatarios utilizar el **icono europeo o un icono o etiqueta equivalente** que cumpla sus especificaciones. El elemento visual principal debe contener el acrónimo **"AI"** en mayúsculas (salvo incompatibilidad con normas nacionales sobre el uso de lenguas), ser legible, reconocible y conservar sus proporciones cuando se redimensione.

El Código propone además una estructura de **dos niveles**:

1. Una señal inmediatamente reconocible que advierta de la intervención de IA.
2. Una segunda capa, recomendada cuando sea técnicamente posible, que indique si el contenido ha sido **generado o modificado** mediante IA y, cuando resulte útil, qué elemento concreto se ha alterado, por ejemplo, el rostro o la voz.

El anexo del Código ofrece tres variantes del icono europeo: "AI Generated", "AI Modified" y una versión básica susceptible de enlazar con información adicional. Las pruebas de usuario realizadas durante su preparación mostraron que las variantes acompañadas de texto explícito se comprendían mejor que un símbolo aislado.

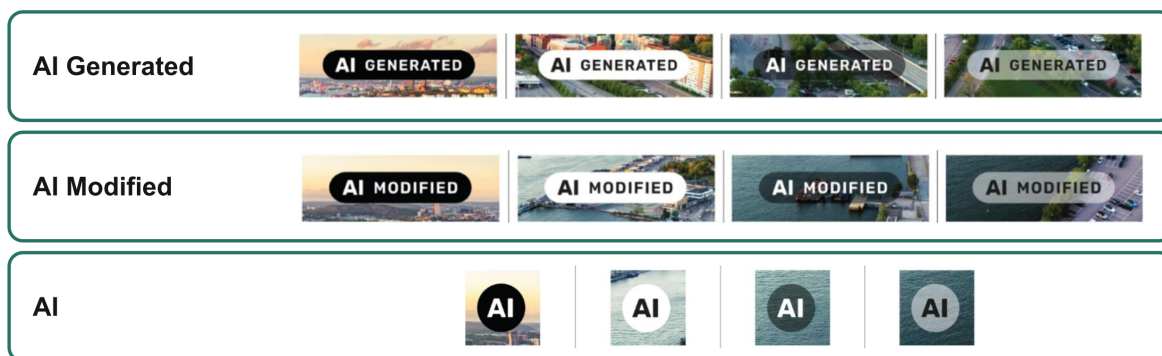


FIGURA 4. Variantes del icono europeo propuestas por el Código de buenas prácticas: AI Generated, AI Modified e icono básico AI. Se conservan los distintivos reproducidos en el artículo, sin rediseñar su identidad.

No existe, fuera de esa vía voluntaria, una frase única impuesta por el Reglamento. La información debe describir fielmente lo ocurrido. "Creado con IA" puede ser excesivo cuando solo se sustituyó una voz; "modificado con IA" puede resultar insuficiente cuando

toda la escena es sintética. La etiqueta revela la intervención artificial, pero no decide quién es jurídicamente autor del contenido.

5.2. Inmediatez, perceptibilidad y duración

La etiqueta **no puede quedar escondida** en los términos y condiciones, tras varias capas de menús o exclusivamente en metadatos que la audiencia no pueda leer. Según el Código, **debe poder reconocerse inmediatamente**, sin exigir una acción específica del usuario ni una atención sostenida, y permanecer visible o audible durante el tiempo suficiente para ser percibida en condiciones normales.

Como regla para sus signatarios, el icono o etiqueta debe incorporarse directamente al contenido, salvo que exista una alternativa equivalente (por ejemplo, una superposición de la interfaz que aparezca materialmente sobre él). También debe conservar suficiente separación respecto de otros elementos y contraste frente al fondo. La "primera exposición" no se refiere a la primera publicación del contenido, sino al momento en que **cada persona** queda expuesta a él por primera vez.

5.3. Soluciones por modalidad y contexto

El Código concreta la presentación de la información del siguiente modo:

- **Imágenes y vídeos:** el icono debe situarse en un lugar apropiado, **sin superposiciones que lo oculten**; el Código menciona como ejemplo la esquina superior derecha. **En vídeo debe aparecer al comienzo** y, cuando sea posible, repetirse a intervalos regulares y, como mínimo, después de interrupciones. Si la ultrafalsificación ocupa solo un fragmento, se recomienda mantener la indicación durante ese segmento.
- **Audio:** cuando no sea posible informar visualmente, debe incluirse **al comienzo de la propia ultrafalsificación un aviso audible breve, perceptible y formulado en lenguaje sencillo**, en la misma lengua del contenido o en inglés. En audios largos, en directo o susceptibles de circular fragmentados, procede añadir recordatorios cuando resulte apropiado y, como mínimo, después de interrupciones. Si existe una pantalla, el aviso audible debe acompañarse de una indicación visual.
- **Textos:** el icono o etiqueta **puede colocarse encima o al comienzo del texto, cerca del titular o en los datos editoriales iniciales**. Cuando solo una parte haya sido generada o manipulada, puede identificarse ese fragmento. Incluso los textos muy breves deben etiquetarse; si hacerlo dentro del propio texto perjudica su legibilidad, el Código permite un aviso contextual adyacente en la interfaz o al inicio de la sesión de exposición.
- **Entornos profesionales cerrados:** la indicación puede **integrarse en la interfaz, el espacio físico u otro medio fácilmente disponible**, siempre que la persona sea informada de forma clara y distinguible antes de quedar expuesta a la ultrafalsificación. Este supuesto hace referencia a un uso exclusivamente interno en un entorno profesional cerrado, como la creación de una ultrafalsificación para un curso interno de ciberseguridad de una empresa.



FIGURA 5. Cuatro formas de presentar la información conforme al artículo 50.5 y al Código de buenas prácticas. Los ejemplos no sustituyen la clasificación previa del contenido ni el análisis de las excepciones.

5.4. El régimen específico de las obras creativas

En obras o programas evidentemente artísticos, creativos, satíricos, ficticios o análogos, el Código desarrolla la **solución atenuada** del artículo 50.4: la ultrafalsificación sigue teniendo que revelarse, pero de una forma que no perjudique la exhibición, el disfrute, la utilidad, la calidad o la explotación normal de la obra.

En medios digitales, el aviso puede situarse fuera pero junto al marco de la imagen o del vídeo, en una descripción, en la interfaz, en los créditos iniciales o finales o mediante un elemento discreto que ofrezca información adicional al pulsarlo o pasar el cursor. No obstante, la existencia del aviso debe poder percibirse sin que la persona tenga que realizar una acción específica. En exposiciones, cines, festivales o soportes físicos, puede facilitarse en la entrada, en el punto de venta, en el programa, en el folleto o en el embalaje.

Esta flexibilidad permite separar dos cuestiones: informar de la intervención de la IA y reconocer la eventual aportación creativa humana. Una fórmula proporcionada puede indicar que una obra identificada con su autor "contiene elementos generados o modificados mediante IA", sin atribuir la autoría al sistema ni borrar al creador.

6. MATRIZ DE SUPUESTOS Y MÉTODO DE CUMPLIMIENTO PARA ORGANIZACIONES Y PROFESIONALES

6.1. Matriz de supuestos: sí, no o depende

La siguiente tabla resume la aplicación de los criterios de las Directrices. Las respuestas presuponen un uso profesional y pueden cambiar si varían el contexto, la audiencia o el proceso editorial.

Tabla 2. Matriz de supuestos y aplicación del artículo 50.4.

Supuesto	¿Etiquetado conforme al artículo 50.4?	Razón principal
Vídeo sintético de una persona semejante a un político pronunciando un discurso inexistente	Sí	Puede constituir una ultrafalsificación y parecer auténtico o verídico, salvo que se trate de un uso estrictamente personal.
Vídeo realista de un consejero delegado sintético dirigido a empleados	Sí	El contexto interno no excluye la obligación; debe informarse antes de la exposición
Imagen de un producto mejorada con IA de manera que altera su aspecto o calidad percibida	Sí, si puede inducir a error	La ultrafalsificación también puede afectar a objetos y publicidad
Anuncio con ratones que hablan como personas	No, en el ejemplo de las Directrices	La escena es manifiestamente irreal y no tiene potencial para parecer auténtica
Actores reales sobre un fondo artificial claramente ficticio	En principio, no	La audiencia no espera que el entorno sea auténtico y los actores no han sido manipulados
Imagen satírica evidente de un político	Sí, con régimen atenuado	La sátira no elimina la obligación, pero permite una divulgación que no perjudique la obra
<i>Influencer</i> sintético mostrando las funciones reales de un producto patrocinado	Sí	La finalidad comercial no convierte por sí sola el contenido en obra creativa o ficticia
Resumen automático de una decisión municipal publicado por un periódico	Sí, salvo excepción editorial	Es texto publicado para informar sobre un asunto de interés público
Alerta meteorológica generada con IA y publicada automáticamente	Sí	Informa al público sobre seguridad; no existe revisión humana
Alerta generada con IA y aprobada sustantivamente por un funcionario competente bajo responsabilidad del organismo	No, si se cumplen ambas condiciones	Revisión humana competente y responsabilidad editorial
Informe para inversores de una sociedad cotizada publicado sin revisión	Sí	La información financiera puede ser materia de interés público
Informe de sostenibilidad revisado por profesionales de cumplimiento y publicado bajo responsabilidad empresarial	No, si la revisión es sustantiva	Puede aplicarse la excepción editorial
Traducción con IA de un artículo sobre política que ha sido revisada antes de publicarse	No, si existe responsabilidad editorial	La traducción final ha sido sometida a revisión humana
Texto revisado por una persona y después reformulado sustancialmente otra vez mediante IA	Sí	La intervención posterior deja sin efecto la revisión de la versión anterior
Asesoramiento regulatorio generado para un único cliente	No	No es un texto publicado al público

Supuesto	¿Etiquetado conforme al artículo 50.4?	Razón principal
Descripción comercial ordinaria de un producto	En principio, no por la rama textual	No trata necesariamente un asunto de interés público; cambiaría si contiene afirmaciones relevantes sobre salud, seguridad o sostenibilidad
Publicación de un creador que obtiene ingresos regularmente y usa IA para informar sobre elecciones sin revisión	Sí	Existe uso profesional y texto público sobre un proceso democrático
Publicación realizada por una persona en una actividad puramente personal y no profesional	No por el artículo 50.4	Opera la exclusión general del artículo 2.10, sin perjuicio de otras normas

6.2. Un método de cumplimiento para organizaciones y profesionales

El Código insiste en que el cumplimiento no debe convertirse en un gesto añadido al final de la publicación. Para sus signatarios, el etiquetado forma parte de un sistema interno de clasificación, diseño, revisión, formación y corrección. Las organizaciones no adheridas no están obligadas a reproducir literalmente ese sistema, pero puede servirles como referencia reconocida para demostrar un cumplimiento efectivo y proporcionado.

Paso 1. Localizar los usos profesionales y asignar funciones

La organización debe conocer qué equipos, agencias y profesionales utilizan sistemas capaces de generar o manipular imagen, audio, vídeo o texto. También debe determinar quién decide el uso y la finalidad, quién actúa como responsable del despliegue, quién clasifica el resultado, quién coloca la etiqueta y quién comprueba que se mantenga al publicarlo o distribuirlo.

Cuando intervengan agencias, profesionales externos o plataformas, los contratos y las instrucciones internas deberían aclarar el reparto de esas tareas. La intervención de un tercero no elimina automáticamente la condición ni la responsabilidad del sujeto bajo cuya autoridad se utiliza el sistema.

Paso 2. Clasificar el resultado antes de publicarlo

Para imagen, audio o vídeo:

1. ¿Existe generación o manipulación mediante IA?
2. ¿El contenido se asemeja apreciablemente a una persona, objeto, lugar, entidad o suceso existente?
3. ¿Puede parecer falsamente auténtico o verídico a la audiencia previsible dentro de ese contexto?
4. ¿Se integra en una obra evidentemente artística, creativa, satírica, ficticia o análoga?

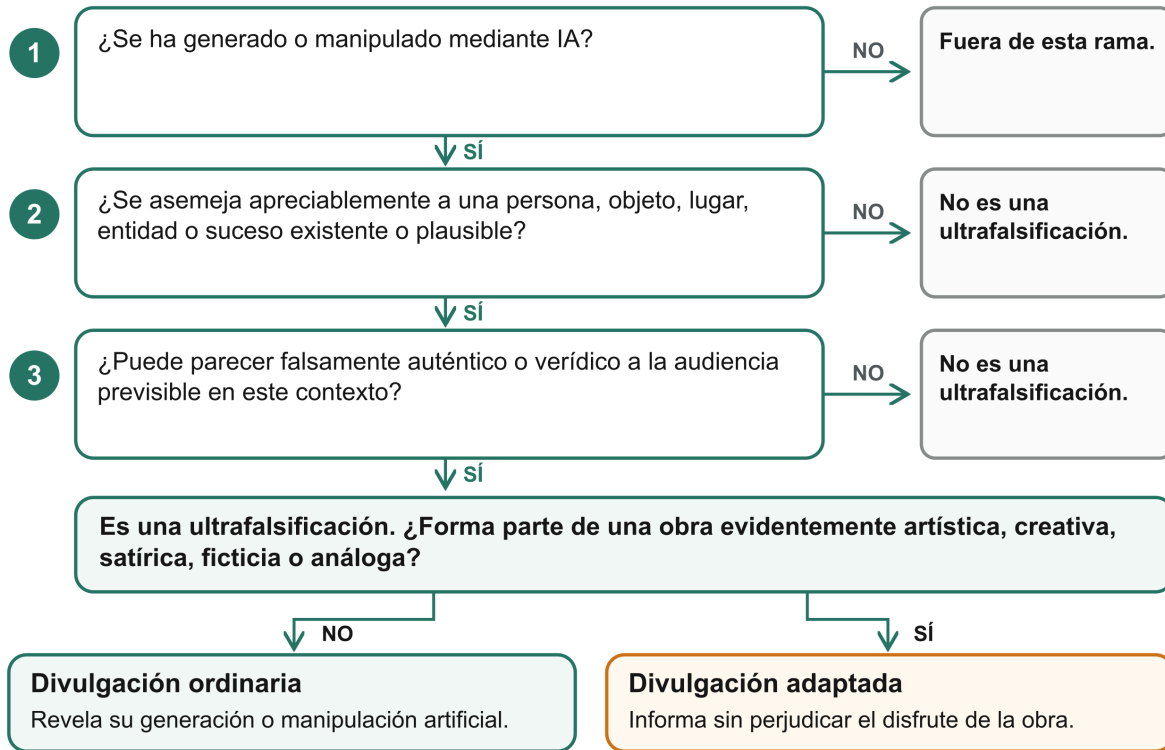
A · Imagen, audio o vídeo

FIGURA 6. Árbol de decisión para imágenes, audios y vídeos, una vez identificados el uso profesional y el responsable del despliegue. El esquema representa el uso ordinario, sin perjuicio de otros supuestos legales.

Para texto:

1. ¿Ha sido generado o manipulado mediante IA?
2. ¿Se va a publicar ante un público indeterminado y suficientemente amplio?
3. ¿Pretende transmitir hechos, opiniones o conocimientos?
4. ¿Trata sobre una materia que merece debate o escrutinio público?
5. ¿La versión final ha sido revisada por una persona competente y existe responsabilidad editorial?

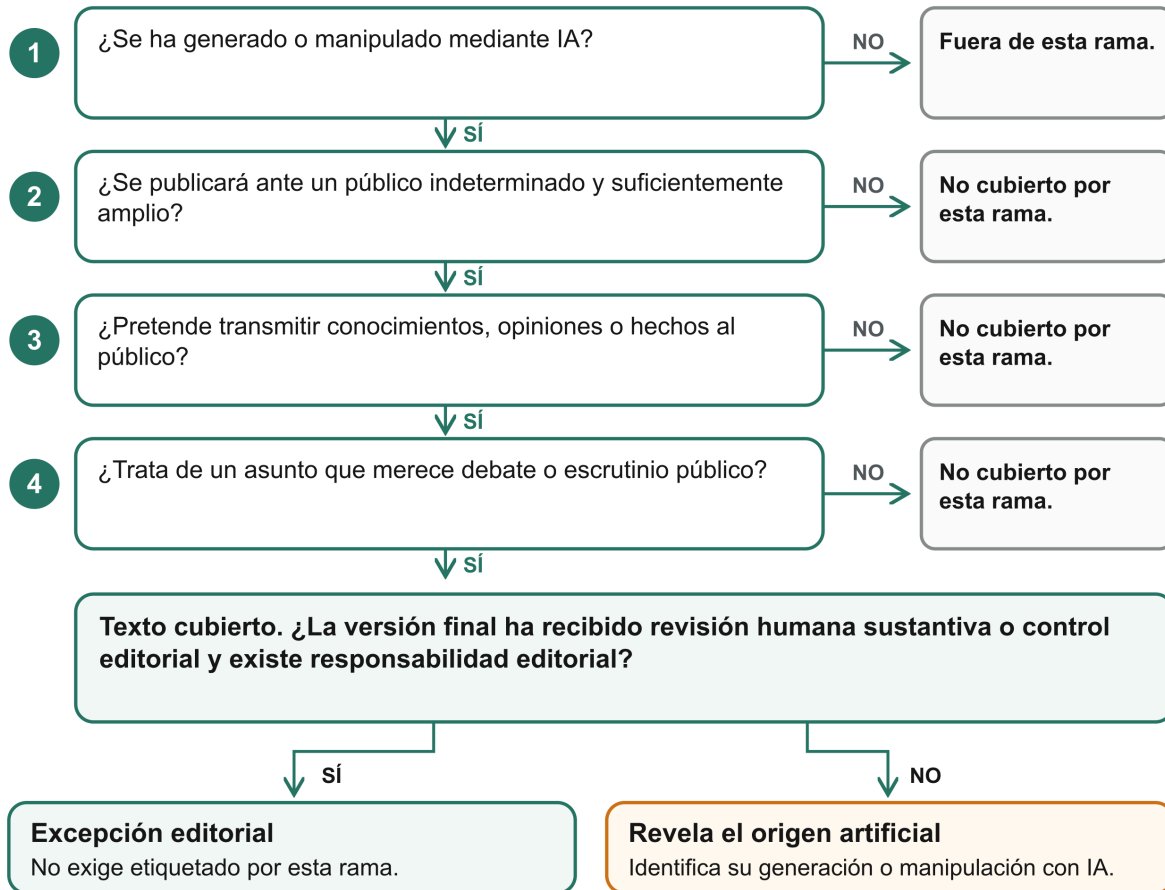
B · Texto

FIGURA 7. Árbol de decisión para textos. La excepción requiere que la revisión humana o el control editorial recaiga sobre la versión final y que una persona física o jurídica asuma la responsabilidad editorial.

Paso 3. Elegir una solución de divulgación y documentar su significado

Quien se adhiera al Código debe optar por el icono europeo o por un icono o etiqueta equivalente conforme a sus especificaciones. La organización debe decidir qué variantes utilizará para distinguir contenido generado y contenido modificado, dónde aparecerán en cada canal y qué información ofrecerá la segunda capa.

El Código anima a publicar una explicación clara del significado de las etiquetas utilizadas, especialmente cuando se emplee un diseño propio o existan variantes contextuales. Esta medida reduce el riesgo de que el público reconozca el símbolo, pero no comprenda qué intervención de IA representa.

Paso 4. Integrar el aviso en el contenido y en su distribución

La etiqueta debe incorporarse en el flujo de publicación: plantilla de imagen, edición de vídeo, reproductor de audio, gestor de contenidos o interfaz correspondiente. Debe comprobarse su perceptibilidad en móvil, escritorio, reproducción iniciada, fragmentos, descargas e interrupciones.

En contenidos creativos se aplicará el régimen específico que evita perjudicar su disfrute y explotación normal. En la distribución posterior conviene coordinarse con plataformas,

editores y distribuidores para preservar el aviso y prever contractualmente quién corrige una pérdida o alteración de la etiqueta.

Paso 5. Regular la revisión humana y la responsabilidad editorial

Si una organización pretende acogerse a la excepción aplicable a determinados textos, debe tratar la revisión como un proceso previo a la publicación y no como una aprobación simbólica. Los prestadores de servicios de comunicación sujetos a estándares editoriales pueden apoyarse, conforme al Código, en sus procedimientos y normas profesionales existentes siempre que satisfagan el artículo 50.4.

Para los demás signatarios, el Compromiso 4 exige establecer o adaptar políticas proporcionadas que identifiquen:

- La persona física o jurídica que asume la responsabilidad editorial, con su nombre, función y datos de contacto.
- Las medidas organizativas y los recursos humanos destinados a efectuar una revisión o control editorial adecuados antes de publicar.

Cuando esa información no sea ya pública, el Código exige publicar los datos de contacto de la función o persona con responsabilidad editorial. No obliga, en cambio, a documentar individualmente cada revisión de cada texto, aunque permite registrar información adicional sobre la naturaleza del control o el grado de intervención de la IA. Todo ello debe aplicarse sin afectar a la libertad de los medios, la independencia editorial ni la protección de las fuentes periodísticas.

Paso 6. Mantener evidencia proporcionada y verificar el funcionamiento

El artículo 50.4 no impone por sí mismo un expediente individual para cada contenido. El Código exige a sus signatarios procesos y documentación interna proporcionados a su tamaño, recursos y capacidades. Esa documentación puede contener una descripción general y ejemplos reales representativos de cómo se aplican las etiquetas y los regímenes especiales.

En organizaciones que utilicen habitualmente IA para crear ultrafalsificaciones o textos cubiertos, el Código exige mantener un proceso que permita verificar que las especificaciones de diseño y ubicación se aplican correctamente y reducir el riesgo de contenidos sin etiquetar o mal etiquetados.

La evidencia puede incluir:

- Política de generación, revisión y publicación de contenidos.
- Criterios de clasificación y ejemplos internos.
- Modelos de iconos, etiquetas y segundas capas por canal.
- Identificación de funciones y responsabilidad editorial.
- Descripción de los perfiles y recursos destinados a la revisión.
- Contratos e instrucciones para agencias y distribuidores.
- Pruebas de interfaz, accesibilidad, duración y conservación de las etiquetas.
- Procedimiento para analizar y corregir avisos ausentes o incorrectos.

Paso 7. Formar, recibir alertas y corregir

El Código exige esfuerzos de sensibilización proporcionados para el personal y los contratistas que aplican o supervisan las medidas. La formación debería cubrir los

supuestos sujetos a divulgación, la excepción editorial, las obras creativas, la accesibilidad y el procedimiento de corrección.

La creación de canales específicos para que usuarios, investigadores, verificadores u otros terceros comuniquen etiquetas ausentes o incorrectas es una recomendación voluntaria del Código. Sin embargo, cuando un signatario recibe una comunicación fundamentada, debe revisar el caso y corregir sin demora indebida los incumplimientos constatados. También debe cooperar con las autoridades competentes conforme al Derecho aplicable.



FIGURA 8. Visión conjunta del método de siete pasos desarrollado en el apartado 6.2. El orden organiza el proceso; la revisión y la corrección deben acompañar también a las modificaciones y redistribuciones posteriores.

7. APLICACIÓN TEMPORAL, SUPERVISIÓN Y SANCIONES

El artículo 50 es aplicable desde el **2 de agosto de 2026**. El periodo transitorio hasta el 2 de diciembre de 2026 introducido para determinados sistemas comercializados con anterioridad afecta únicamente al mercado técnico del artículo 50.2. No aplaza la obligación de divulgación del artículo 50.4.

Las Directrices precisan el tratamiento del contenido anterior a esa fecha:

- Las ultrafalsificaciones generadas o manipuladas antes del 2 de agosto de 2026 no tienen que etiquetarse retroactivamente.
- Los textos generados o manipulados y publicados antes de esa fecha tampoco.
- Si un texto fue generado o manipulado antes, pero se publica a partir del 2 de agosto de 2026, sí debe etiquetarse cuando concurren los requisitos.
- La Comisión anima a identificar contenidos anteriores cuando sea posible, sin exigir esfuerzos desproporcionados como auditar íntegramente bases históricas o modificar embalajes ya impresos.

Las autoridades nacionales de vigilancia del mercado ejercerán con carácter principal la supervisión. Las personas y organizaciones pueden presentar reclamaciones al amparo del artículo 85.

El incumplimiento puede dar lugar, conforme al artículo 99, apartado 4, letra g), a multas administrativas de hasta **15 millones de euros** o, cuando el infractor sea una empresa, de hasta el **3 % de su volumen de negocios mundial total correspondiente al ejercicio financiero anterior**, aplicándose el importe superior. Para las pymes, incluidas las empresas emergentes, cada multa queda limitada al menor de los dos importes correspondientes. El artículo 99, apartado 6 bis, extiende esta regla del importe inferior a las pequeñas empresas de mediana capitalización (*small mid-cap enterprises* o SMC) respecto de las multas previstas para el incumplimiento del artículo 50. La decisión debe atender a las circunstancias concretas, como la naturaleza, la gravedad y la duración de la infracción, sus consecuencias, su carácter intencional o negligente y el grado de cooperación, entre otros factores.

La adhesión al Código de buenas prácticas no constituye una prueba concluyente de cumplimiento, pero ofrece un marco reconocido para demostrarlo. Quien no se adhiera puede utilizar medios alternativos adecuados, aunque deberá estar preparado para explicar con mayor detalle cómo satisfacen el Reglamento y las Directrices.

8. LÍMITES DEL ARTÍCULO 50.4

Como se ha señalado al principio del artículo, estos deberes de transparencia pretenden informar a las personas de manera más eficaz sobre el origen artificial o la manipulación del contenido.

Esta concreción reduce la fragmentación, pero no elimina todas las tensiones.

- **Informar del origen no equivale a certificar la verdad o su legalidad:** Una ultrafalsificación correctamente etiquetada puede seguir siendo falsa, difamatoria o ilícita; un texto señalado como generado mediante IA puede contener información exacta. Igualmente, una ultrafalsificación bien etiquetada puede resultar ilícita conforme a otras normas, como las de propiedad intelectual, protección de datos o protección del derecho al honor y a la propia imagen.
- **Uniformidad frente a precisión informativa:** El icono común facilita el reconocimiento de una misma señal en distintos sectores y Estados miembros. Sin embargo, un símbolo genérico "AI" puede no explicar si el resultado fue íntegramente generado, parcialmente modificado o simplemente asistido por una herramienta. El Código intenta resolver esa tensión mediante las variantes "AI Generated" y "AI Modified" y una segunda capa que puede precisar qué elemento se ha alterado.
- **Libertad artística, autoría y explotación normal de la obra:** Informar de que una obra contiene elementos generados o modificados mediante IA no atribuye la condición de autor al sistema ni niega por sí mismo la eventual aportación creativa de una persona. No obstante, una etiqueta imprecisa como "creado por IA" podría invisibilizar esa contribución.
- **Responsabilidad editorial sin vigilancia permanente:** El Compromiso 4 del Código de Buenas Prácticas intenta hacer verificable la excepción editorial: exige a determinados signatarios identificar la responsabilidad editorial, describir los recursos destinados a la revisión y publicar un contacto cuando no exista ya información pública. El equilibrio es razonable, pero deja una cuestión probatoria. Una política general puede demostrar que existe un sistema de control, aunque no siempre permitirá acreditar qué revisión recibió una publicación concreta. Exigir un expediente exhaustivo para cada texto

reforzaría la prueba, pero también podría generar cargas desproporcionadas e interferir en los procesos editoriales.

- **Riesgo de fatiga y de confianza inversa:** Un uso excesivo, ambiguo o poco visible de las etiquetas puede reducir su eficacia. Si el público encuentra constantemente advertencias genéricas, puede terminar ignorándolas.
- **La exclusión de determinados usos personales puede generar una brecha:** El contenido creado inicialmente en un contexto personal puede alcanzar posteriormente una difusión masiva. Una ultrafalsificación publicada por una persona particular puede ser reproducida, descontextualizada o viralizada por terceros, aunque su creador inicial no estuviera sujeto a la obligación, generando desinformación.

9. CONCLUSIONES

El artículo 50.4 nace como solución a posibles problemas derivados de la creación de contenido con sistemas de IA, como el fraude, la manipulación, el engaño o la desinformación. En esa búsqueda de transparencia, el artículo 50.4 genera otro tipo de tensiones, como la exclusión de determinados contenidos de uso personal, la posible fatiga derivada de la generalidad de las etiquetas y las tensiones relacionadas con el derecho de autoría y la distribución de las obras.

En el artículo se recogen los pasos que se deben seguir y los aspectos que se tienen que tener en cuenta para aplicar correctamente el precepto. No todo el contenido generado por IA conlleva la obligación de etiquetar. Estos pasos se pueden sintetizar en las siguientes siete preguntas:

1. ¿Quién decide utilizar el sistema y con qué finalidad?
2. ¿Actúa dentro de una actividad profesional?
3. ¿Qué parte del contenido ha sido generada o manipulada mediante IA?
4. ¿El contenido es una ultrafalsificación o un texto público sobre un asunto de interés público?
5. ¿Existe un régimen atenuado o una excepción realmente aplicable?
6. ¿La información será clara, distinguible y perceptible en la primera exposición de cada persona?
7. ¿Puede demostrarse que el proceso funciona también al revisar, publicar y distribuir el contenido?

Para los responsables del despliegue la introducción de este tipo de etiquetas resulta molesta. Esto se debe principalmente a que la creación de contenido es un ámbito en el que se premia la inmediatez y el etiquetado supone una carga adicional. Además, es una realidad que existe una mala percepción social del contenido generado con IA. La sociedad tiende a menospreciar y a no valorar de igual manera este tipo de contenido.

Para solventar estas tensiones de la mejor manera posible, las empresas y los profesionales deberán actuar conforme a los principios de proporcionalidad y responsabilidad proactiva. El objetivo de los responsables será utilizar un método de etiquetado que permita a la vez cumplir las obligaciones del artículo 50.4; no perjudicar la percepción de la autoría ni la experiencia del consumo del contenido; y que no suponga un esfuerzo desproporcionado.

Es imprescindible que cada responsable protocolice la creación y etiquetado de este tipo de contenido, más aún si supone una práctica habitual. El protocolo deberá poder asegurar

el cumplimiento de las obligaciones y poder demostrar dicho cumplimiento, nuevamente, de manera proporcional, sin sobredocumentación ni esfuerzos desproporcionados.

REFERENCIAS

Selección de fuentes primarias y orientaciones institucionales utilizadas para preparar este análisis.

Última comprobación: 27 de agosto de 2026.

- [1] [Texto consolidado del Reglamento \(UE\) 2024/1689, Reglamento de Inteligencia Artificial, a 27 de julio de 2026](#). Véanse especialmente los considerandos 132 a 137 y los artículos 2, 3.4, 3.60, 50, 85, 99 y 113.

<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A02024R1689-20260727>

- [2] [Reglamento \(UE\) 2026/1744, Digital Omnibus sobre IA](#). Introduce, entre otras modificaciones, el periodo transitorio limitado del artículo 50.2 y el tratamiento de las pequeñas empresas de mediana capitalización en el régimen sancionador.

<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32026R1744>

- [3] [Corrección de errores del Reglamento \(UE\) 2024/1689, de 9 de octubre de 2025](#). Corrección de la denominación española «ultrafalsificación» en el considerando 134, el artículo 3.60 y el artículo 50.4.

https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202590802

- [4] [Comisión Europea: Directrices sobre la aplicación de las obligaciones de transparencia de determinados sistemas de IA conforme al artículo 50](#). C(2026) 5054 final, 20 de julio de 2026; véanse especialmente los apartados 9, 12 a 24, 111 a 144 y 145 a 155.

<https://ec.europa.eu/newsroom/dae/redirection/document/131215>

- [5] [Comisión Europea: preguntas frecuentes sobre las obligaciones de transparencia del artículo 50](#). Aclaraciones institucionales sobre el alcance de las obligaciones de transparencia.

<https://digital-strategy.ec.europa.eu/es/faqs/transparency-obligations-under-article-50-ai-act>

- [6] [AESIA: Preguntas y respuestas sobre el Reglamento de IA: artículo 50](#). Síntesis institucional española de las obligaciones de transparencia previstas en el artículo 50 y de sus objetivos.

<https://aesia.digital.gob.es/es/actualidad/recursos/ria-articulo-50>

- [7] [Código de buenas prácticas sobre transparencia de los contenidos generados por IA](#). Especialmente su sección 2 sobre responsables del despliegue.

<https://ec.europa.eu/newsroom/dae/redirection/document/129555>

- [8] [Comisión Europea: evaluación del Código de buenas prácticas](#). La adhesión facilita la demostración del cumplimiento, pero no constituye prueba concluyente.

<https://digital-strategy.ec.europa.eu/en/library/commission-opinion-assessment-code-practice-transparency-ai-generated-content>

- [9] [Comisión Europea: iconos europeos para etiquetar contenidos generados por IA](#). Información institucional sobre los iconos comunes propuestos para el etiquetado.

<https://digital-strategy.ec.europa.eu/en/policies/eu-icons-labelling-ai-generated-content>