Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

EU AI Act Article 50: Transparency and AI Content Marking

> Reliable Context is a kernel of truth: a system of record for organizations that fuse private and public information and must defend every value they act on. One canonical record per real thing, a cited source and an observation date behind every value, unknowns declared rather than left blank, and any past state of the record reproducible.

Published by Reliable Context · 2026-10-29 · ai act article 50 · ai act transparency obligations · ai generated content labeling · ai act watermarking · ai act chatbot disclosure

Original article: https://reliablecontext.com/articles/eu-ai-act-article-50-transparency-marking-synthetic-content



In this article

1. What Article 50 regulates, in one paragraph
2. Two roles: provider and deployer
3. Paragraph 1: disclosure when people interact with AI
4. Paragraph 2: machine-readable marking of synthetic output
5. Paragraph 3: emotion recognition and biometric categorisation
6. Paragraph 4: deepfakes and AI-generated text on public-interest matters
7. Paragraph 5: how and when the information is given
8. Paragraphs 6 and 7: other law, and codes of practice
9. The Code of Practice and the Commission guidelines
10. What the Code asks for technically
11. Technical options and what each one proves
12. Worked example: one company, three roles
13. Relation to Article 13 for high-risk systems
14. Enforcement and penalties
15. Checklist
16. Sources

What Article 50 regulates, in one paragraph

Article 50 of the EU Artificial Intelligence Act (Regulation (EU) 2024/1689) is the Act's transparency rule for AI systems that are not necessarily high-risk. It imposes four separate duties: providers must make chatbots and other directly interactive systems disclose that they are AI (paragraph 1); providers of generative systems must mark outputs in a machine-readable way that makes them detectable as artificial (paragraph 2); deployers of emotion recognition and biometric categorisation systems must inform the people exposed (paragraph 3); and deployers must disclose deepfakes and certain AI-generated text published on matters of public interest (paragraph 4). Paragraph 5 says how and when the information is given, paragraph 6 keeps other rules intact, and paragraph 7 provides for codes of practice. These obligations apply from 2 August 2026 under Article 113 of the Act, a date the Digital Omnibus on AI did not move, apart from a short transition for existing generative systems described below.

This article reads each paragraph against the official text, separates provider from deployer duties, and examines what "marked in a machine-readable format and detectable" means in practice, and what a mark does and does not prove.

Two roles: provider and deployer

Each duty belongs to a role. The Act defines a provider (Article 3(3)) as a person or body that develops an AI system or general-purpose AI model, or has one developed, and places it on the market or puts it into service under its own name or trademark, whether for payment or free of charge. A deployer (Article 3(4)) is a person or body using an AI system under its authority, except in a personal non-professional activity.

One organisation can hold both roles: a bank that puts a customer-service assistant built on a third-party model into service under its own name is its provider and its deployer.

Paragraph	Who carries the duty	What must happen	Main exceptions in the text
50(1)	Provider	People informed they are interacting with AI	AI involvement obvious; law-enforcement uses
50(2)	Provider (including of general-purpose AI systems)	Outputs marked machine-readably and detectable as artificial	Standard editing; no substantial alteration of input; law enforcement
50(3)	Deployer	Inform people exposed to emotion recognition or biometric categorisation	Law-enforcement uses
50(4), first subparagraph	Deployer	Disclose deepfakes	Law enforcement; lighter duty for evidently artistic or satirical works
50(4), second subparagraph	Deployer	Disclose AI text published on public-interest matters	Law enforcement; human review plus editorial responsibility
50(5)	Both	Clear, distinguishable, accessible, at first exposure	None

All rows paraphrase Article 50 of Regulation (EU) 2024/1689; the Commission's AI Act Service Desk displays the same article as amended.

Paragraph 1: disclosure when people interact with AI

Paragraph 1 is a design obligation on the provider. The system must be "designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system" (Article 50(1)). The exception for cases where AI involvement is obvious uses a reasonable-person test: obvious "from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use". Recital 132 adds that the characteristics of vulnerable groups, by age or disability, should be taken into account where the system is intended to interact with them.

The obviousness exception is a judgment worth writing down: a voice agent on a telephone line is rarely obvious to callers, while a labelled "AI assistant" panel in a software tool may be.

Paragraph 2: machine-readable marking of synthetic output

Paragraph 2 is the provision most often summarised as "watermarking", which undersells it. The text requires providers of AI systems, "including general-purpose AI systems, generating synthetic audio, image, video or text content", to "ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated" (Article 50(2)). It then sets four quality attributes: the technical solutions must be "effective, interoperable, robust and reliable as far as this is technically feasible", taking into account the type of content, implementation costs and the generally acknowledged state of the art.

Three terms deserve precise reading:

- **Marked** means a signal is attached to or embedded in the output. The Act does not prescribe a technique. Recital 133 lists candidates: "watermarks, metadata identifications, cryptographic methods for proving provenance and authenticity of content, logging methods, fingerprints or other techniques".
- **Machine-readable** means software, not only a person, can read the signal.
- **Detectable** means someone other than the provider must be able to establish, from the output, that it was generated or manipulated. A mark nobody can check does not satisfy the second half of the sentence.

The exceptions are narrow. Marking is not required "to the extent the AI systems perform an assistive function for standard editing or do not substantially alter the input data provided by the deployer or the semantics thereof" (Article 50(2)). Spell-checking a human-written paragraph is the paradigm case; generating the paragraph from a two-line prompt is not.

Transitional period. The Digital Omnibus on AI, Regulation (EU) 2026/1744 (published in the Official Journal on 24 July 2026 and in force since 27 July 2026, the third day after publication), added a paragraph to Article 111: providers of generative systems "that have been placed on the market before 2 August 2026 shall take the necessary steps in order to comply with Article 50(2) by 2 December 2026". Systems placed on the market on or after 2 August 2026 have no such period. The transition covers paragraph 2 only.

Paragraph 3: emotion recognition and biometric categorisation

Paragraph 3 obliges deployers of an emotion recognition system or a biometric categorisation system to "inform the natural persons exposed thereto of the operation of the system" and to process personal data under the GDPR, Regulation (EU) 2018/1725 and Directive (EU) 2016/680 as applicable (Article 50(3)). Other provisions of the Act prohibit or classify as high-risk some uses of these systems; paragraph 3 governs only disclosure.

Paragraph 4: deepfakes and AI-generated text on public-interest matters

Paragraph 4 has two subparagraphs with different triggers.

Deepfakes. The Act defines a deep fake (Article 3(60)) 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". Text is not part of this definition. Deployers of a system that produces such content "shall disclose that the content has been artificially generated or manipulated" (Article 50(4)). For "evidently artistic, creative, satirical, fictional or analogous" works, the duty is reduced to disclosing the existence of generated or manipulated content in a way that "does not hamper the display or enjoyment of the work".

Public-interest text. Deployers of a system that generates or manipulates text "published with the purpose of informing the public on matters of public interest" must disclose that it is AI-generated or manipulated. The duty does not apply "where the AI-generated content has undergone a process of human review or editorial control

and where a natural or legal person holds editorial responsibility for the publication" (Article 50(4)). Both conditions are needed: review alone, without someone accountable for the publication, does not meet the exception.

Relying on the exception turns an editorial process into a compliance fact. The Code of Practice (discussed below) asks signatories outside regulated media to maintain policies that identify the person with editorial responsibility and describe the organisational measures for review, while stating that this "does not entail having to document individual instances of human review".

Paragraph 5: how and when the information is given

Every disclosure under paragraphs 1 to 4 must be given "in a clear and distinguishable manner at the latest at the time of the first interaction or exposure" and must "conform to the applicable accessibility requirements" (Article 50(5)). A disclosure buried in terms of service does not meet it.

Paragraphs 6 and 7: other law, and codes of practice

Paragraph 6 states that paragraphs 1 to 4 do not affect the high-risk requirements of Chapter III and are without prejudice to other Union or national transparency obligations. Recital 137 adds that compliance with these transparency duties "should not be interpreted as indicating that the use of the AI system or its output is lawful". A correctly labelled deepfake can still be defamatory or infringe personality rights.

Paragraph 7 was replaced by the Digital Omnibus. As amended, the Commission "shall encourage and facilitate the drawing up of codes of practice" on "the detection, marking and labelling of artificially generated or manipulated content", and, taking utmost account of the AI Board's opinion, "shall assess whether adherence to those codes of practice is adequate to ensure compliance with the obligations laid down in paragraphs 2 and 4"; if a code is inadequate, the Commission may adopt an implementing act with common rules (Regulation (EU) 2026/1744, Article 1). The original text had instead empowered the Commission to approve codes by implementing act.

The Code of Practice and the Commission guidelines

Two Commission instruments now sit beside the article.

The Code of Practice on Transparency of AI-generated Content. According to the Commission's Code of Practice page, the final code was published on 10 June 2026 after three drafting rounds (first draft 17 December 2025, second draft 3 March 2026). It has two sections: Section 1 for providers (marking and detection) and Section 2 for deployers (labelling of deepfakes and AI-generated text). The page states that "adherence to the code is voluntary" while the Article 50 requirements "are legal obligations", that the Commission and the AI Board have confirmed the code is an adequate voluntary tool to demonstrate compliance, that providers and deployers who comply by other means "will have to demonstrate that those measures are adequate", and that about 190 organisations had signed by the end of July 2026. It is a separate instrument from the General-Purpose AI Code of Practice for general-purpose AI model providers.

The guidelines on Article 50. The Commission's guidelines page (last updated 6 August 2026) describes guidelines that clarify who is a provider or deployer for each duty, define concepts such as directly interactive systems, synthetic content, deepfakes and public-interest text, and give in-scope and out-of-scope examples, including the standard-editing exception. It names the enforcers: national market surveillance authorities, the AI Office for systems under its supervision, and the European Data Protection Supervisor for EU institutions.

What the Code asks for technically

The Code of Practice (PDF) is the most concrete reading of paragraph 2 available. Its key positions for providers:

1. **No single technique suffices.** Signatories acknowledge that "in most cases, in particular for content that can be disseminated online, no single marking technique suffices to meet the four requirements in Article 50(2)".
2. **At least two layers.** Until a single technique can be shown to meet all four requirements, signatories implement "a multi-layered marking approach" with "at least two layers of machine-readable marking": digitally signed metadata (Sub-measure 1.1.1) and an imperceptible watermark (Sub-measure 1.1.2).
3. **Signed metadata.** Where the format supports metadata, signatories record whether content is AI-generated or manipulated, and "all recorded information will be digitally signed and time-stamped ... in a secure and tamper-evident manner".
4. **Text.** Because "free-form text cannot transport metadata", a watermark alone is considered sufficient for that content type; for free-form text longer than 200 tokens watermarking still applies, even with lower reliability.
5. **Fingerprinting or logging are optional supplements.** "Relying on fingerprinting or logging alone is not considered sufficient".
6. **Detection must be available to others.** Signatories offer a detection solution (a public specification, software, or an API) to deployers, end-users and other legitimate parties, in principle free of charge.
7. **Richer provenance is encouraged.** Measure 1.3 encourages recording the AI system name, provider name, a timestamp and optionally the model identifier and version.

For deployers, Section 2 uses the EU "AI" icon or an equivalent label.

Technical options and what each one proves

The techniques are not interchangeable. Each binds a different claim to the content, and each fails differently.

Technique	What it carries	Typical failure	What a positive check proves
Signed metadata (e.g. a C2PA manifest)	Signed statements about origin and edits, hash-bound to the content	Stripped by re-encoding, screenshots, some uploads	The signer made these statements about this exact content
Invisible watermark	A signal embedded in pixels, samples or token choices	Weakened by cropping, compression, paraphrase	Content probably passed through that system

Technique	What it carries	Typical failure	What a positive check proves
Fingerprint	A compact content description matched against a database	Needs a stored record; near-duplicates collide	Content resembles something recorded earlier
Provider logging	Provider-side record of outputs	Privacy cost; only the provider can query	The provider's records show this output

The C2PA specification defines a claim as "a digitally signed and tamper-evident data structure that references a set of assertions", distinguishes hard bindings (cryptographic hashes over the asset) from soft bindings such as fingerprints and invisible watermarks, and notes that an asset "can become separated from its C2PA Manifest due to removal or corruption of asset metadata", in which case a fingerprint or watermark can be used to look the manifest up again (C2PA 2.5). The specification also states that it should not provide value judgments about whether a given set of provenance data is good or bad: it validates structure and integrity, not truth. The Content Authenticity Initiative describes the same mechanism under the name Content Credentials.

That last point matters for anyone using marks as evidence. A valid mark establishes **origin**: this output came from that system, or was signed by that party. It does not establish that the content is accurate, that the person depicted consented, or that the text's claims are supported. Article 50 is a disclosure rule; it is not an accuracy rule. Accuracy is a separate property, established by tracing each statement to a source, which is what provenance and attribution records are for.

Worked example: one company, three roles

Halden Therapeutics AG is a fictional company used for illustration.

Use	Halden's role	Paragraphs	Evidence kept
Patient-information chatbot on a licensed model, under Halden's name	Provider and deployer	50(1) disclosure in the first message; 50(2) for generated text	Obviousness assessment; vendor's marking and detection documentation
Vendor image generator; one campaign image shows a realistic laboratory that does not exist	Deployer (vendor is provider)	50(4) deepfake label at first exposure	Label design and placement; test that CMS upload keeps the vendor's signed metadata
AI-drafted drug-shortage updates, reviewed by a named editor	Deployer	50(4) text exception	Policy naming the person with editorial responsibility and the review process

Relation to Article 13 for high-risk systems

Article 13 applies only to high-risk systems and faces deployers, not the public: high-risk systems must be "sufficiently transparent to enable deployers to interpret a system's output and use it appropriately" and come with instructions for use (Article 13(1)-(2)). A high-risk system that also generates content or talks to people is

subject to both (Article 50(6)). After the Digital Omnibus, Chapter III Sections 1 to 3, which include Article 13, apply from 2 December 2027 for Annex III systems and 2 August 2028 for Annex I systems; Article 50 already applies.

Enforcement and penalties

Non-compliance with "transparency obligations for providers and deployers pursuant to Article 50" falls in the tier of administrative fines of up to EUR 15 000 000 or, for an undertaking, up to 3 % of total worldwide annual turnover for the preceding financial year, whichever is higher (Article 99(4)(g)). For SMEs, including start-ups, the lower of the two amounts applies (Article 99(6)). Penalty rules are laid down by Member States.

Checklist

- [ ] Inventory systems that converse, generate content, or perform emotion recognition or biometric categorisation; record your role for each.
- [ ] Paragraph 1: document the obviousness judgment or design the disclosure into the first interaction.
- [ ] Paragraph 2: document marking layers, detector and third-party access; record reasons for any standard-editing exception; for systems on the market before 2 August 2026, plan for 2 December 2026.
- [ ] Paragraph 4: define which outputs are deepfakes under Article 3(60); fix label design and placement; name the person with editorial responsibility if relying on the text exception.
- [ ] Paragraph 5: verify disclosures are clear, given at first exposure, and accessible.
- [ ] Test that your own pipelines preserve upstream marks.
- [ ] Sign the Code of Practice or document an alternative of equivalent adequacy.

This article explains EU law as of 28 October 2026 for educational purposes and is not legal advice. Reliable Context keeps a cited, dated source behind every value it holds, which addresses the accuracy question that marking leaves open.

Sources

1. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L, 12 July 2024 (EUR-Lex ELI) — Articles 3(3), 3(4), 3(60), 13, 50, 99, 113; recitals 132-137.
2. Regulation (EU) 2026/1744 (Digital Omnibus on AI), Official Journal L, 24 July 2026 (EUR-Lex ELI) — replaced Article 50(7); Article 50(2) transition to 2 December 2026; new high-risk application dates.
3. European Commission AI Act Service Desk, Article 50 — consolidated display of Article 50.
4. European Commission AI Act Service Desk, Article 13 — transparency for high-risk systems.
5. European Commission, Code of Practice on Transparency of AI-generated Content (policy page) — status, adequacy confirmation, timeline, signatory count.
6. European Commission, Code of Practice on Transparency of AI-generated Content (PDF), June 2026 — multi-layer marking, signed metadata, watermarking, detection, editorial-responsibility policies.
7. European Commission, Guidelines on transparency obligations under Article 50 (landing page) — scope of the guidelines and enforcing authorities.

8. C2PA Technical Specification 2.5, October 2026 — claims, assertions, hard and soft bindings, separation of manifests.
9. Content Authenticity Initiative, How it works — Content Credentials overview.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://eur-lex.europa.eu/eli/reg/2024/1689/oj
2. https://eur-lex.europa.eu/eli/reg/2026/1744/oj
3. https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-50
4. https://digital-strategy.ec.europa.eu/en/policies/code-practice-ai-generated-content
5. https://digital-strategy.ec.europa.eu/en/policies/guidelines-transparency-ai-generated-content
6. https://ec.europa.eu/newsroom/dae/redirection/document/129555
7. https://spec.c2pa.org/specifications/specifications/2.5/specs/C2PA_Specification.html
8. https://contentauthenticity.org/how-it-works
9. https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-13

THE PUBLISHER

About Reliable Context

Reliable Context is a system of record for organizations that fuse private and public information and must defend every value they act on.

The guarantees

One canonical record per real thing: duplicates resolve by rule, ambiguity is shown rather than hidden, and related organizations relate without ever being merged. A verbatim, cited source and an observation date sit behind every value. Unknowns are declared as not found, blocked or not applicable with a reason, never left as an empty cell. Any past state of the record is reproducible, together with what changed since.

Where it runs

It runs inside the customer's perimeter, one database per organization. Permitted sources, model routing, reader clearance and retention are set as the customer's own policies, and every value is exportable with its chain of custody for an auditor or regulator.

What it is not

It is not a vector database or a RAG framework, not a data catalog, MDM or data-quality suite, not a CRM or data-enrichment vendor, and not a chatbot or AI assistant. Where a reader needs one of those, the honest answer is to say so and point them there.

Visit

See Reliable Context.

Reliable Context is in production use in two sectors; more than six thousand of its identity decisions were made by rule and none has been reversed. Articles published here explain general practice, cite their primary and academic sources, and name the jurisdiction of any regulatory rule they describe. They are not legal, regulatory or compliance advice for a specific organization, and requirements change.

Website: https://reliablecontext.com

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.