AI Technology Partners

ENTERPRISE AI MANAGED SERVICES

The CMO's guide to AI content disclosure

Regulatory requirements and ethical best practices for enterprise publishers

AUTHOR

John Treadway
Chief Executive Officer & Co-Founder, AI Technology Partners, Inc.

DATE

August 27, 2026
Version 2.0

DISCLOSURES

AI-Assisted Content

This guide was researched and drafted with substantial assistance from an AI language model (Claude, by Anthropic), including web-based research, source synthesis, and initial drafting. It was reviewed and edited by the author before publication. The author discloses this AI involvement as a matter of policy and good practice; on these facts (substantive human review, editorial responsibility, and no deceptive representation), no general disclosure duty appears to require it for a document of this kind. Readers should independently verify any fact, figure, date, or characterization of law before relying on it.

Not Legal Advice

This guide is provided for general informational and educational purposes only. It is not legal advice and does not create an attorney-client relationship between the reader, the author, or AI Technology Partners, Inc. AI disclosure regulation is evolving rapidly, interpretations vary by jurisdiction, and the application of any rule depends on specific facts. Before adopting or changing disclosure practices, enterprises should review their approach with qualified legal counsel in each jurisdiction where their content is published or consumed.

Currency of Information

This guide reflects publicly available information as of August 27, 2026. AI regulation, enforcement priorities, and industry standards in this area are changing quickly. Readers should confirm current requirements before acting, particularly for the EU AI Act as amended by Regulation (EU) 2026/1744, the Commission's evolving guidance and Code of Practice, FTC enforcement activity, and state-level statutes.

Table of Contents

1. Executive Summary

Marketing organizations now operate under a patchwork of binding regulation, active enforcement, and voluntary standards governing when AI involvement in content must be disclosed. The three most consequential developments for enterprise publishers, plus a fourth practical shift worth planning around now:

1. **The EU AI Act's transparency obligations (Article 50) took effect August 2, 2026.** Many enterprises using third-party AI tools are "deployers" under the Act (the same company can be a provider for systems it offers under its own brand), and the key duty for text content is narrower than commonly reported: labeling is required only for AI-generated text published to inform the public on matters of public interest, and content that has undergone genuine human review or editorial control is exempt. Deepfake-style image, audio, and video content carries a separate, broader disclosure duty. Fines reach EUR 15 million or 3% of worldwide turnover, whichever is higher.
2. **The US has no single federal AI disclosure statute.** The FTC applies its existing deception authority (FTC Act Section 5), the Endorsement Guides (16 CFR Part 255), and the 2024 Consumer Reviews and Testimonials Rule to AI-created content with full force. The FTC is actively enforcing the reviews rule, civil penalties for knowing violations run to $53,088 each, and state laws (New York, California) add further requirements for synthetic performers and digital replicas.
3. **Voluntary frameworks are converging on a risk-based, materiality-driven standard.** The EU's Code of Practice on Transparency of AI-Generated Content (June 2026) and IAB's AI Transparency & Disclosure Framework V2 (August 2026) both point in the same direction: disclose when AI involvement would change what a reasonable audience believes about the content, and maintain documented human oversight when you do not disclose.
4. **Plan for AI text to become detectable.** Major providers are embedding machine-readable marks in text output under the EU Code of Practice: Anthropic now marks output from supported Claude models worldwide (older models are still being updated), and Google has watermarked Gemini text since 2024. About 190 organizations had signed the Code by end-July, 82 of them, including OpenAI, Microsoft, and Meta, on the provider-side marking section. Public detection tools still trail the marks, so treat detectability as a forward-looking risk rather than a present capability: verbatim or lightly edited AI text published today may be identifiable later.

Aug 2, 2026	**€15M / 3%**	**$53,088**	**82 / 152**
EU AI Act Article 50 in force	Maximum EU fine (or global turnover)	Max FTC penalty per knowing rule violation	EU Code signatories: provider-side / deployer-side

The practical posture for most enterprises: build a documented editorial review process (which addresses the EU text obligation), disclose AI involvement wherever content depicts people, voices, events, or experiences that could be mistaken for real (which addresses deepfake rules and FTC deception risk), and adopt a consistent internal decision framework so disclosure choices are deliberate rather than ad hoc.

2. First, Know Your Role: Provider vs. Deployer

The EU AI Act assigns different obligations depending on your role. Getting this right determines most of what follows.

Role	Who it is	Core obligation
Provider	Develops an AI system (or has it developed) and places it on the market or puts it into service under its own name or trademark. Model and tool vendors such as OpenAI, Anthropic, Google, and Adobe; a brand offering an AI experience under its own name can qualify too.	Embed machine-readable marks in synthetic outputs (Article 50(2)); ensure interactive AI informs users they are dealing with AI (Article 50(1)).
Deployer	Uses an AI system under its own authority in a professional context. Most enterprises producing blog posts, campaign imagery, video, or social content with third-party AI tools.	Label covered AI text on public-interest topics and disclose deepfakes (Article 50(4)); inform people exposed to emotion-recognition or biometric-categorization systems (Article 50(3)).

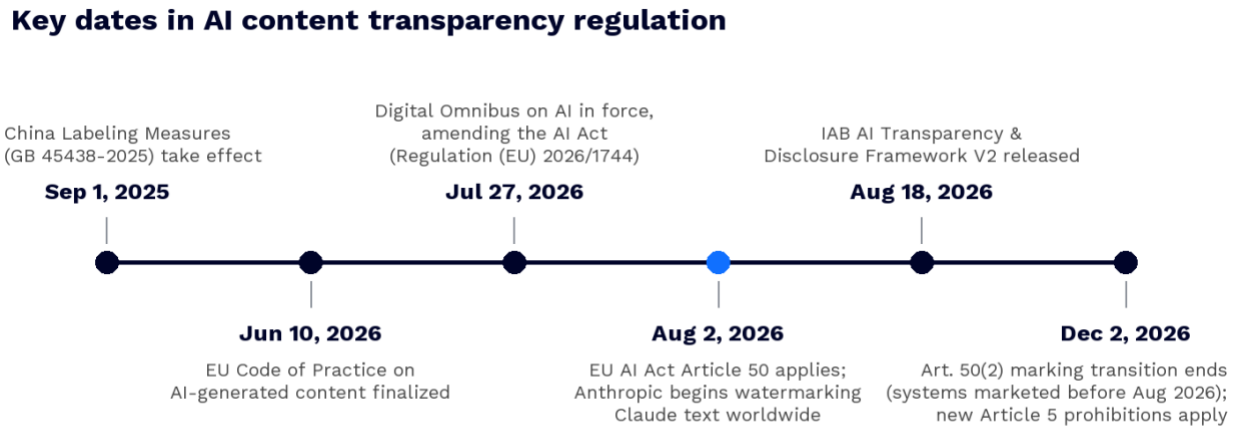
Individual employees acting under company instruction (writers, designers, agencies working on your behalf) are not separate deployers; the legal entity is. Two caveats worth checking with counsel:

- If your enterprise fine-tunes a model, white-labels an AI tool, or offers an AI-powered experience to customers under your own brand, you may take on provider obligations for that system. The test is Article 3(3), placing a system on the market or into service under your own name or trademark; fine-tuning alone does not decide it.
- Territorial scope reaches beyond your headquarters. Non-EU companies are in scope when their AI-system outputs are used in the EU, and EU audience reach is a strong signal rather than the whole test: role, system placement, and how outputs are used all factor in. A US enterprise publishing content consumed by EU audiences should assume exposure and assess rather than dismiss.

Source: European Commission, Transparency obligations under Article 50 of the AI Act (FAQ)

3. The EU AI Act: What Article 50 Requires of Enterprise Publishers

Article 50 of Regulation (EU) 2024/1689 has applied since **August 2, 2026**. The Act was amended just before that date: Regulation (EU) 2026/1744, the Digital Omnibus on AI, was published July 24, 2026 and entered into force July 27, 2026, and it left Article 50's substantive obligations and the August 2 date in place, so work from the consolidated text. One transition matters here: providers of synthetic-content systems already on the market before August 2, 2026 have until December 2, 2026 to meet the Article 50(2) marking duty. Systems placed on the market from August 2 get no transition, and deployer duties never had one. Content generated before August 2, 2026 does not require retroactive labeling, though the Commission encourages it.



Key regulatory and industry dates affecting enterprise AI content, 2025–2026.

3.1 AI-Generated Text (Article 50(4), second subparagraph)

This is the provision most relevant to blogs, thought leadership, and editorial content. Deployers must label AI-generated or manipulated text **only when three cumulative conditions are met**:

#	Condition	What it means
1	Published	The text is made available to the public.
2	Informs the public	It is published to inform, not merely to sell or serve internal purposes.
3	Matter of public interest	Politics, public administration, justice, fundamental rights, public security, public health, environmental protection, consumer safety, or economic, financial, scientific, or cultural developments subject to public debate.

That last category is broad. Enterprise thought leadership on regulation, market economics, scientific developments, or public health can fall within it, though not every article in those domains will: purpose and the link to public debate matter. Routine product marketing, sales collateral, and most B2B content will often sit outside "informing the public" in the civic sense the Commission's

guidelines describe, but there is no categorical marketing or B2B exemption. This is a judgment call, not a bright line. Document the rationale for anything categorized as out of scope, including its purpose and the public-debate link.

The human review exemption

Text that has undergone qualifying human review or editorial control does not need to be labeled. The Act sets no AI-percentage threshold: the exemption applies however much AI was involved in drafting, provided the review is genuinely substantive, and heavier AI involvement raises the practical bar for demonstrating that it was. (Provider-side marking of raw output is a separate duty and continues regardless.) The Commission's guidance sets a substantive bar:

- **Human review** means deliberate examination of the substance of the content by a person with relevant knowledge and professional judgment.
- **Editorial control** means a responsible editorial entity (for example, an editor-in-chief or equivalent) with authority to approve, alter, or reject the content on substantive grounds, including fact-checking and source verification.
- Spell-checking, grammar correction, and other superficial or procedural checks **do not** qualify.
- Someone must hold ultimate legal responsibility for publication (editorial responsibility).

A deployer relying on the exemption should be ready to demonstrate it: retain evidence of the substantive review, the editorial control, and who held editorial responsibility. No particular form of record is mandated. In practice the documented editorial process, more than any label, is the compliance artifact that matters for text content.

3.2 Deepfakes: Image, Audio, and Video (Article 50(4), first subparagraph)

Deployers must disclose that content is AI-generated or manipulated when it constitutes a "deepfake" under Article 3(60): AI-generated or manipulated image, audio, or video that resembles existing (or plausibly existing) persons, objects, places, entities, or events and would falsely appear authentic or truthful.

Key points from the Commission's guidance:

- Disclosure must reach the viewer **at first exposure at the latest**, in a clear, distinguishable, perceivable form (visible or audible labels). Machine-readable watermarks embedded by the tool vendor do **not** satisfy the deployer's disclosure duty on their own.
- Context matters. If the intended audience would not expect the content to be authentic (standard VFX, obviously stylized creative), it may not meet the "false appearance" test. Photorealistic AI imagery of people, places, or events presented in a way audiences could take as real generally does, provided it resembles an existing or plausibly existing subject; photorealism alone does not decide it.
- For evidently artistic, creative, or satirical works, disclosure is still required but may be done in a way that does not hamper enjoyment of the work.

For marketing teams, this is the provision with the widest practical reach: photorealistic AI imagery in campaigns, synthetic voiceovers resembling real speech, AI-generated "customer" scenes, and

digitally altered footage of real people or places all warrant scrutiny. None of these is automatically a deepfake in every context; each can be.

3.3 Interactive AI (Article 50(1))

Article 50(1) places this duty on the provider of the system: interactive AI must be designed and developed so that people know they are interacting with AI from the start of the first interaction, unless that is obvious to a reasonably well-informed person (an exception the Commission instructs be interpreted restrictively). The catch for brands: a company that puts a chatbot, agent, or avatar into service under its own name or trademark is usually the provider of that system under Article 3(3). So the practical instruction stands: identify the AI at the start of the conversation. Where a vendor genuinely remains the provider, verify contractually that its implementation delivers the notice.

3.4 Enforcement and Penalties

National market surveillance authorities enforce Article 50, with the AI Office covering certain GPAI-based and platform-integrated systems. Article 50 infringements fall under Article 99(4)(g): fines reach **EUR 15 million or 3% of total worldwide annual turnover, whichever is higher**. SMEs and start-ups face whichever ceiling is lower, and the 2026 Omnibus amendment added a similar capped tier for small mid-caps.

3.5 New Article 5 Prohibitions (December 2, 2026)

The Omnibus amendment also added two prohibited practices aimed directly at synthetic media. From December 2, 2026, Article 5 bans placing on the market, putting into service, or using AI systems to generate or manipulate realistic intimate imagery of an identifiable person without that person's explicit consent, or to generate child sexual abuse material. These sit in the Act's highest penalty tier: up to EUR 35 million or 7% of worldwide turnover. No legitimate marketing program goes near this conduct, but the prohibitions still matter operationally, because they also reach tools whose design makes such output a reasonably foreseeable outcome absent adequate safeguards. That makes tool selection and acceptable-use policies the practical control point for anyone running generative pipelines at scale.

Note the date: December 2, 2026 now carries two distinct obligations. The Article 50(2) marking transition for pre-existing systems ends, and these Article 5 prohibitions begin.

Sources

- Article 50, EU AI Act (official text)
- Consolidated AI Act text as amended by Regulation (EU) 2026/1744 (CELEX 02024R1689-20260727)
- European Commission FAQ on Article 50
- Commission Guidelines on Transparency of AI-Generated Content
- Quick Facts: Transparency rules for AI systems

4. United States: FTC Authority and State Laws

4.1 The Federal Framework

There is no standalone general-purpose federal AI disclosure statute, though sectoral and content-specific federal laws can apply. The operative framework is:

- **FTC Act Section 5** (deceptive acts and practices). Undisclosed AI content violates Section 5 when it creates a materially deceptive net impression: when consumers would reasonably believe they are seeing real people, real experiences, or unaltered products. There is no categorical rule against undisclosed AI use as such.
- **Endorsement Guides, 16 CFR Part 255** (2023 revision). Endorsements must reflect honest opinions of real endorsers, material connections must be disclosed clearly and conspicuously, and these principles apply to AI-generated and AI-modified endorsement content with the same force as human-created content. The Guides add no blanket AI-use disclosure duty; AI-specific notice turns on whether omitting it would deceive.
- **Consumer Reviews and Testimonials Rule** (2024). Fake reviews and testimonials, explicitly including AI-generated reviews not based on genuine product experience, are prohibited outright. Disclosure does not cure a fabricated review.

Practical implications commonly drawn from FTC positions and enforcement (secondary sources vary in how they characterize these; validate specifics with counsel):

Disclosure type	What it covers	Satisfied by
Commercial relationship	The existing endorsement/sponsorship disclosure ("#ad").	Clear, proximate sponsorship label.
AI involvement	Whether omitting AI's role (identity, experience, alteration) creates a deceptive net impression.	A clear AI-use disclosure where needed to prevent deception. One label does not do the other's job.

- **Make disclosures clear and conspicuous for the duty and the medium.** The reviews rule defines "unavoidable" for the disclosures it covers; endorsement guidance stresses proximity and format-appropriate prominence. In practice: near the content rather than buried in footnotes or hashtag clusters, on-screen early for video, visible without expansion for static posts.
- **Synthetic personas cannot pass as real people** in commercial content where that impression would deceive; disclose the synthetic identity when needed to avoid it. AI-generated testimonials cannot be substantiated by personal experience, because there is none.
- **Judge tooling by materiality and deception.** Routine edits (grammar checking, analytics, cosmetic retouching, lighting and color adjustments) typically fall below the disclosure line; substantive manipulation (face swaps, voice cloning, AI lip-sync, fabricated scenes and results) typically sits above it. Treat the lists as shorthand for one question: does the net impression mislead?

- **Liability is distributed.** Brands bear primary exposure, with agencies and creators carrying independent liability. Contracts allocating liability between brand and creator do not bind the FTC.
- Civil penalties run to **$53,088 per violation** for knowing violations of the reviews rule and other specified provisions, and multiple acts can be alleged as separate violations. Treat the figure as a ceiling for covered rule violations rather than an automatic consequence of any undisclosed AI use.

4.2 State Laws

- **New York (A8887-B, effective June 9, 2026)**: conspicuous disclosure is required where an advertiser knows a commercial ad features a synthetic performer. Audio-only ads, AI translation of a real performer's speech, and certain promotions for expressive works are exempt.
- **California AB 1836**: consent-based digital replica protections for deceased personalities in specified expressive uses; it governs consent and likeness rather than AI disclosure.
- **California AB 2602** (operative January 2025): contract enforceability requirements for digital replicas of living performers.
- Additional states are active. Colorado's current framework (SB26-189 plus Attorney General rulemaking) covers automated decision-making that materially influences consequential decisions, along with chatbot safety; it does not regulate AI advertising disclosure. The state layer is moving quickly and warrants periodic legal review.

Sources

- FTC Endorsement Guides, 16 CFR Part 255
- FTC Rule on Consumer Reviews and Testimonials
- Dynamis LLP, AI Disclosure in 2026: Developments and Practical Steps
- Promise Legal, FTC AI Disclosure Rules for Creators (2026)

5. China: A Highly Prescriptive Labeling Regime (For Global Enterprises)

Enterprises publishing into China face a highly prescriptive, multi-actor framework. The Measures for Labeling AI-Generated Synthetic Content, with mandatory national standard GB 45438-2025, took effect **September 1, 2025**. Duties attach by actor, with a different trigger for each:

- **Generation providers.** Explicit labels (visible text, icons, watermarks, captions, or audio cues suited to the modality) are required in specified scenarios, and implicit metadata labels enabling identification and traceability are required more broadly. Where explicit labels belong in the file, they must survive download, copy, and export. One exception: under Article 9, a provider may supply unlabeled output at a user's request, with an agreement fixing the user's labeling responsibilities and with logs kept.

- **Propagation platforms.** Platforms must check metadata, add or reinforce notices, and handle unlabeled uploads based on user declarations or their own detection.
- **Publishing users.** Anyone publishing AI-generated content through those platforms must declare it and use the platform's labeling functions. For this guide's readers, this is the duty that bites: a marketing team posting AI-generated creative to a Chinese platform holds it, whatever the tool vendor or the platform does.

Regime	Jurisdiction	Mechanism	Interoperable with others?
EU Article 50(2)	European Union	Machine-readable watermark, provider-side	Basis for C2PA alignment; not identical to China's schema
China GB 45438-2025	China (distribution into)	Explicit visible label + implicit metadata	No; distinct schema, do not assume equivalence
C2PA / Content Credentials	Global, voluntary	Cryptographically signed provenance manifest	Supports EU goals; does not satisfy China Measures alone

Note for compliance architecture: GB 45438's metadata schema is not identical to C2PA and does not map directly onto the EU's machine-readable marking requirement. Compliance with one regime does not automatically satisfy the others. Extraterritorial reach is not fully settled; confirm applicability with counsel before building or excluding controls.

Sources

- Cyberspace Administration of China, Measures for Labeling of AI-Generated Synthetic Content (official text)
- Covington, China Releases New Labeling Requirements for AI-Generated Content
- Bird & Bird, New AI Content Labelling Rules in China vs. the EU AI Act

6. Voluntary Frameworks and Ethical Standards

6.1 EU Code of Practice on Transparency of AI-Generated Content (June 2026)

A voluntary, Commission-facilitated code that providers and deployers can sign to demonstrate compliance with the marking and labeling obligations of Article 50(2), (4), and (5). It has been assessed as adequate by the Commission and the AI Board, so signatories gain legal certainty and predictability. Non-signatories must demonstrate compliance through alternative means and may face more information requests. It also sets design, placement, and accessibility expectations for AI labels, and official EU icons are now available for optional use.

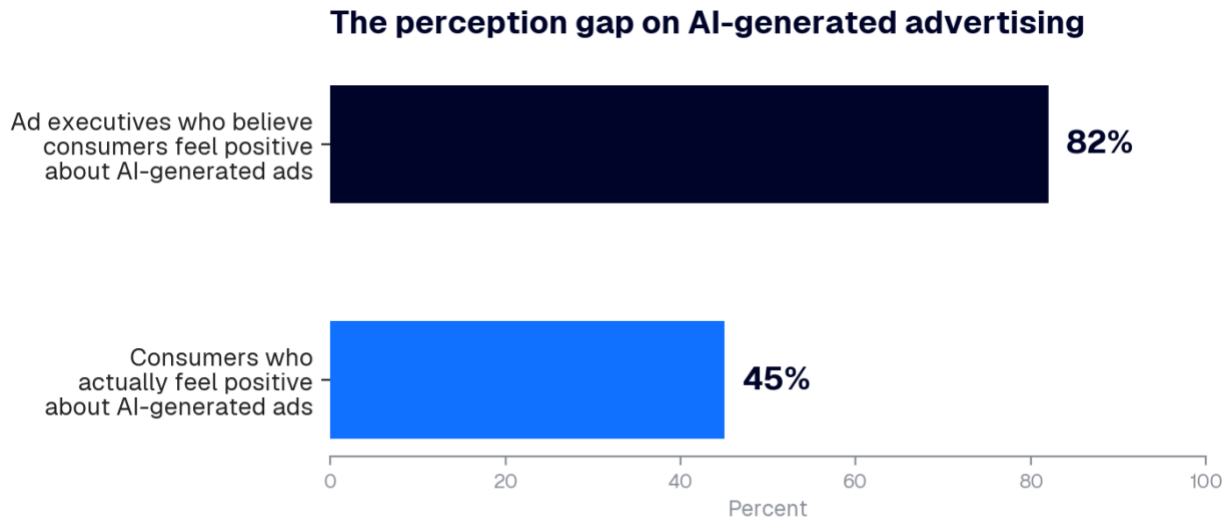
Recommendation: enterprises with meaningful EU exposure and significant synthetic content volume should evaluate signing, recognizing that whether it pays off depends on role, scope, systems, and governance cost; those with lighter exposure should at minimum align internal label design to the Code's conventions.

Source: Code of Practice on Transparency of AI-generated Content

6.2 IAB AI Transparency & Disclosure Framework V2 (August 2026)

The advertising industry's self-regulatory reference point. V2 translates a risk-based, materiality-driven approach into practical protocols for when and how AI involvement should be disclosed across consumer-facing content: AI-generated and AI-assisted text, imagery, video, audio, synthetic voices, digital twins, and AI-powered consumer interactions. It is industry guidance: it explicitly defers to state, federal, and international regulation wherever they conflict.

The core ethical logic, which most published frameworks now share: **disclosure is warranted when AI involvement is material to how a reasonable audience would evaluate the content**, particularly where content depicts people, experiences, or events. Assistive uses that do not materially change what the audience sees (research, outlining, grammar, formatting) generally do not warrant disclosure. This is the practical convergence point, with one caveat: the EU and China rules described above add categorical triggers that apply regardless of materiality.



Source: IAB research with Sonata Insights, 2026

For the second consecutive IAB/Sonata study, ad executives overestimated how positively consumers view AI-generated advertising; the gap widened from 32 points in 2024 to 37 in 2026, an argument for testing clear, consistent disclosure rather than betting on silence.

Source: IAB AI Transparency & Disclosure Standards V2

6.3 Content Provenance: C2PA / Content Credentials

C2PA is the open technical standard for cryptographically signed provenance metadata: how a file was made, with what tools, and whether AI was involved (creator identity is optional). It is voluntary, and it increasingly functions as connective tissue across regimes: a control that supports the EU's machine-readable marking goals and produces tamper-evident, audit-grade documentation, though credentials record how a file was made rather than whether its content is true, and adoption alone satisfies neither Article 50 nor China's Measures. Adobe, Microsoft, and major camera and platform vendors ship support today; not every product or platform preserves or surfaces credentials yet. Enterprises building durable content operations should adopt Content

Credentials in the creative pipeline now, pair them with generation logs, and test what survives their toolchain, treating provenance as infrastructure rather than a compliance checkbox.

Source: C2PA Specifications

6.4 Text Watermarking: Marks First, Detectors Later

A practical development that changes the disclosure calculus. In August 2026, Anthropic began embedding machine-readable watermarks in Claude text output worldwide, on supported models launched from August 2, 2026 (older models are still being updated), stating that the marking implements its Article 50(2) commitments under the EU Code of Practice. Google has watermarked Gemini text with SynthID since 2024 and reports at-scale testing with minimal quality or speed effects. Coverage is spreading, but it is not universal: 82 organizations, including OpenAI, Microsoft, and Meta, signed the Code's provider-side marking section by end-July 2026, and a signature alone does not tell you which models, versions, and modalities carry marks today. Verify by provider and model rather than assuming.

How it works, in brief: the design Google documents, and a common approach generally, biases the model's probabilistic word choices using a keyed pseudorandom source, leaving a statistical signal a matching detector can read given enough text (Anthropic describes an embedded machine-readable mark without publishing its mechanism). Vendors report no meaningful change to quality or meaning. The mark travels with copied text and can persist through light editing; it weakens as the model's wording is replaced, and heavy rewriting, paraphrasing, or translation will usually defeat it, though none of this is a guaranteed relationship, and very short passages may carry no reliable signal at all. Generated image files from supported tools carry C2PA signed metadata instead, which format conversion or screenshots can strip.

Four implications for enterprise content teams:

1. **Treat detectability as a risk posture, in both directions in time.** Public tooling trails the marks today: Anthropic has published no general text detector for Claude (detection documentation is promised), and Google's public verification covers image, video, and audio rather than text. Take no comfort from the gap. Marks embedded now can be read by whatever tooling ships later, against archives that do not go away, and running content through third-party detectors of uneven quality has already become a public sport around executive bylines. Text published undisclosed today may be flagged, fairly or unfairly, years from now.
2. **Keep the legal test separate from the watermark.** The Article 50(4) exemption turns on substantive human review or editorial control plus editorial responsibility. Reviewed and approved AI text can qualify while still carrying a mark, and heavy editing that happens to clear a mark earns nothing legally without genuine review. A substantive human rewrite will often do both at once; treat that as a convenience, and document the review either way.
3. **Never strip watermarks or provenance metadata.** Hold this as an absolute rule of enterprise policy, whatever a given jurisdiction requires: deliberate removal can turn a defensible judgment call into evidence of intent, and the content may be flagged by other means anyway.

4. **Do not treat detector output as proof of misconduct.** A detected mark means the model processed the content, not that it authored it. Translation, AI-applied proofreading, summarization, and quoted AI text can all leave marks on substantially human work.

Action on AI-drafted text	Watermark status	EU 50(4) status
Publish verbatim, no review	Mark present on supported models; readable by matching tools	Labeling required if the public-interest text test is met
Light edit, keep AI phrasing	Mark weakened; may still be detectable	Exemption not established; risk remains
Substantive human rewrite	Mark usually removed; not guaranteed	Exemption applies if review is genuine and documented
Translate or proofread human-written text	Mark may appear on assistive pass	Outside scope; assistive use, no label needed

Sources

- Anthropic, How Claude Marks AI-Generated Content (support article)
- Google DeepMind, Watermarking AI-Generated Text and Video with SynthID
- Euronews, EU compliance delivered globally: Anthropic to watermark Claude's output worldwide
- Search Engine Journal, Anthropic To Mark Claude Text & Files Under EU AI Act Code
- Zvi Mowshowitz, AI Text Watermarking Is Free And Good (Aug 21, 2026)

6.5 Emerging Ethical Consensus

Across regulators, industry bodies, and published commentary, five principles recur:

1. **Materiality over blanket labeling.** Label what could mislead; do not dilute disclosure by stamping everything. This principle operates inside the binding EU and China rules, never in place of them.
2. **Human accountability.** A named human or editorial function should hold responsibility for anything published, whatever the drafting method; where the EU text exemption is in play, that responsibility is the compliance mechanism itself. "Human-in-the-loop" review before publication remains the anchor practice; as agentic workflows automate publishing, governance shifts to defined approval gates and audit trails.
3. **Honesty about depiction.** Content that shows people, voices, experiences, or results the audience would take as real warrants the highest scrutiny everywhere, and usually disclosure.
4. **Consistency and documentation.** Inconsistent, ad hoc disclosure decisions raise compliance and trust risk alike. A written policy, applied uniformly, supports the compliance case and the trust story, though it is no defense by itself. Survey evidence points the same direction: YouGov's 17-market study and Washington State University's 2024 survey both found majorities favoring transparency about AI use in marketing (results vary by country and use

case), and some brands (Le Creuset, Aerie) now market their human production processes as a trust position.

5. **No AI posing as human in engagement.** An observed pattern, reported by writers and newsletter authors rather than measured by anyone: AI bots posting expert-sounding comments, building rapport with an author over a long thread, then pivoting to a pitch. The judgment in this guide is that this kind of undisclosed AI engagement does outsized damage to trust because the deception is personal. AI-driven comments, replies, and outreach that simulate human interest belong on the prohibited list as enterprise policy, in every jurisdiction and ahead of any statute requiring it.

7. Decision Framework: What to Do, When

The table below reflects a combined reading of EU, US, and voluntary standards. "Disclose" means a clear, proximate, human-readable label; "Document" means retain internal records of process and rationale.

Scenario	EU AI Act obligation	US exposure	Recommended action
Blog/thought leadership, AI-drafted, substantively human-reviewed	Exempt via human review / editorial control, with editorial responsibility	Low, absent deceptive claims	No label required. Document the review. Optional voluntary disclosure as a trust position.
AI-generated article, no substantive review, public-debate topic	50(4) label required if published to inform the public	Low to moderate	Label clearly, or better, add genuine editorial review. Do not rely on the public-interest gray zone.
Fully automated content at scale (SEO pages, product copy)	Assess "informing the public" on the specific facts	Deception risk if claims are false	Assess with counsel; substantiate all claims; consider labeling.
Photorealistic AI imagery of people, places, or events	Deepfake disclosure if it could appear authentic	Section 5 deception risk	Disclose visibly at first exposure. Add Content Credentials as supporting evidence (they do not replace the visible label).
Synthetic voice or AI video resembling real speech/footage	Deepfake disclosure likely required	High if used in ads/endorsements	Disclose visibly/audibly. Obtain consent for any real person's likeness; check NY/CA statutes (NY exempts audio-only ads).
AI persona/avatar in sponsored or endorsement content	Deepfake disclosure; if interactive, 50(1) notice (the brand is often the provider)	High: Endorsement Guides; reviews rule where testimonial-like	Disclose the commercial relationship, and the synthetic identity where needed to avoid deception. Never present it as a real customer.
AI-generated reviews or testimonials, no real experience	N/A (prohibited conduct)	Prohibited outright (2024 Rule)	Do not publish. Disclosure does not cure fabrication.

Scenario	EU AI Act obligation	US exposure	Recommended action
Obviously stylized/artistic AI creative	Deepfake test decides; artistic style alone is no exemption	Low	If it meets the deepfake definition, disclose in a way that does not hamper the work; if not, document the assessment.
Brand chatbot or AI agent interacting with customers	50(1): inform at first interaction unless obvious. A provider duty; brands offering a bot under their own name are usually the provider	FTC expects transparency	Identify the AI at conversation start; interpret "obvious" narrowly. If a vendor is the provider, verify notice contractually.
AI for research, outlines, translation, grammar only	Assistive; outside labeling scope where output is not materially altered	Minimal	No disclosure needed, but note some assistive uses may still leave a watermark signal.
Publishing verbatim/lightly edited AI text, undisclosed	Depends on content type above	Detection risk compounds exposure	Treat it as potentially identifiable, now or later. Add genuine, documented review, or disclose. Never strip marks.
Content distributed into China	China Measures apply to distribution into China; duties attach by actor	N/A	Declare AI content and use platform labels as the publishing user; confirm provider and platform duties under GB 45438.

8. Building the Program: Ten Actions for the Next 90 Days

1. **Inventory AI use in the content supply chain.** Every tool, every workflow, every agency and freelancer. You cannot make disclosure decisions about uses you have not mapped. Include agency work: the enterprise remains the deployer when contractors operate under its responsibility and control, while an agency acting under its own authority may hold a role of its own.
2. **Adopt a written AI content disclosure policy.** Define the assistive/substantive line, the scenarios requiring labels, approved label language and placement, and who decides edge cases. Uniform application matters more than any single threshold choice.
3. **Formalize and document editorial review.** For EU text exposure, this is the compliance mechanism. Require substantive review by someone with subject-matter judgment, with authority to alter or reject, and keep records: reviewer identity, nature of changes, sign-off, date. No particular form of record is mandated; this set is simply easy to defend.
4. **Set hard rules for the highest-risk conduct.** Make these enterprise policy, deliberately stricter than any single law. Any content showing people, voices, places, events, experiences, or results that audiences could take as real gets a visible disclosure and a consent check. Add

three explicit prohibitions: no stripping of watermarks or provenance metadata, no AI systems claiming to be human or simulating human interest, and no fabricated testimonials or reviews under any framing.

5. **Standardize label language and placement.** Short, plain, proximate: "AI-generated image," "Created with AI," "This is a virtual persona." Align to the EU Code of Practice conventions and IAB V2 protocols. Ensure accessibility.
6. **Push obligations into vendor and agency contracts.** Seek role-specific terms. From AI tool vendors: representations on Article 50(2) marking where they will give them, and evidence access, audit, and cure rights where they will not. From agencies and creators: disclosure of material AI use, adherence to your labeling policy, and indemnity. Scope disclosure clauses to material use; blanket everything-clauses collide with the nuanced legal tests. Contracts do not shift FTC liability off the brand.
7. **Deploy provenance tooling and plan for ambient detection.** Turn on Content Credentials (C2PA) in the creative stack, retain generation logs, and keep a provider-by-provider inventory of which models, versions, and modalities mark their output. Do not brief teams that all major-model text is watermarked by default; it is not, yet.
8. **Separate your jurisdiction logic.** EU duties attach by role: provider marking and interaction notices, deployer deepfake disclosure and the public-interest text label. FTC exposure turns on deception and commercial relationships. China assigns actor-specific labeling duties triggered by distribution into China. Encode three distinct tests, not one merged one.
9. **Train the teams that touch content.** Writers, designers, social, PR, agencies, and the executives whose bylines appear on thought leadership. The premise of this action, drawn from experience rather than any enforcement dataset: the typical failure starts with someone who never knew the line existed.
10. **Assign ownership and review cadence.** Name an accountable owner (marketing operations or brand governance, with legal partnership), audit quarterly, and track moving parts: the EU common icon, FTC guidance updates, state statutes, and platform-specific rules (YouTube, Meta, TikTok, LinkedIn).

9. A Note on the Strategic Choice

The regulatory floor is lower than much of the commentary suggests: a well-run editorial process can exempt most enterprise text in the EU, and honest content that deceives no one carries limited FTC deception risk, though endorsement, review, and substantiation rules still apply on their own terms. The open question for each brand is whether to stop at the floor.

There is a defensible case on both sides. Blanket disclosure may dilute meaning and, in some audiences, discount otherwise strong work; that effect is worth testing with your audience rather than assuming. The working hypothesis this guide favors, offered as judgment rather than measured result, is that selective, materiality-based disclosure paired with visible human

accountability earns more trust than silence. Survey evidence is at least directionally supportive: majorities in YouGov's 17-market study and Washington State University's 2024 survey want transparency about AI use in marketing. Some brands are going further and marketing their human review process itself.

Watermarking and improving detection tilt this calculation. The question is no longer only "should we disclose" but "are we comfortable with this content being identified as AI-generated by someone else, on their timing." The prudent planning assumption is that involuntary exposure costs more than voluntary disclosure, and that detection can run retroactively: what passes today may be flagged in two years, accurately or not, by better tools against an archive that does not go away. Partial AI use carries its own version of the risk: AI-written passages inside otherwise human work are the kind of thing readers and classifiers pick out, and the plausible worst case lands where the content's message depends on authenticity or effort. None of this is measured. It is, in this guide's judgment, the direction the incentives point.

The right answer depends on brand positioning, audience, and category. The wrong answer is not deciding, and letting individual teams improvise.

10. References

EU (primary sources)

- Regulation (EU) 2024/1689 (AI Act), consolidated text as amended (CELEX 02024R1689-20260727)
- Regulation (EU) 2026/1744 (Digital Omnibus on AI), OJ L, July 24, 2026
- European Commission, FAQ: Transparency obligations under Article 50 of the AI Act
- European Commission, Guidelines on Transparency of AI-Generated Content
- European Commission, Code of Practice on Transparency of AI-generated Content
- European Commission, Strong Backing for the Code of Practice on Transparency of AI-Generated Content (signatory announcement, July 31, 2026)
- European Commission, Quick Facts: Transparency rules for AI systems
- AI Act Explorer (unofficial consolidated text and analysis)

United States

- FTC Endorsement Guides, 16 CFR Part 255
- FTC, Rule on the Use of Consumer Reviews and Testimonials
- FTC, Guides Concerning the Use of Endorsements and Testimonials (business guidance)
- New York A8887-B (synthetic performer disclosure, amending GBL § 396-b)
- Colorado SB26-189, Automated Decision-Making Technology
- Colorado Attorney General, AI rulemaking (ADMT and chatbot safety)
- Governor of California, AB 2602 and AB 1836 signing announcement
- FTC, Inflation-Adjusted Civil Penalty Amounts for 2025
- OMB Memorandum M-26-11 (2026 civil penalty levels held at 2025 amounts)
- Dynamis LLP, AI Disclosure in 2026
- Promise Legal, FTC AI Disclosure Rules for Creators (2026)

China

- Cyberspace Administration of China, Measures for Labeling of AI-Generated Synthetic Content (official text)
- SAMR / SAC, GB 45438-2025, Labeling Method for Content Generated by Artificial Intelligence (mandatory national standard)
- Covington & Burling (Inside Privacy), China Releases New Labeling Requirements for AI-Generated Content
- Bird & Bird, New AI Content Labelling Rules in China vs. the EU AI Act

Industry and technical standards

- IAB, AI Transparency & Disclosure Standards V2
- C2PA (Coalition for Content Provenance and Authenticity), Specifications
- Cloud Security Alliance, Research Note on EU AI Act Article 50

Watermarking and detection

- Anthropic, How Claude Marks AI-Generated Content (support article)
- Google, Making It Easier to Understand How Content Was Created and Edited (May 2026)
- Google DeepMind, Watermarking AI-Generated Text and Video with SynthID
- Euronews, Anthropic to watermark Claude's output worldwide (Aug 11, 2026)
- Search Engine Journal, Anthropic To Mark Claude Text & Files Under EU AI Act Code
- Forbes, Claude Will Now Leave A Watermark On Everything It Writes (Aug 13, 2026)
- Zvi Mowshowitz, AI Text Watermarking Is Free And Good (Aug 21, 2026)
- Scott Aaronson, watermarking research background (via OpenAI)

Secondary sources cited above reflect practitioner interpretation as of August 2026 and should be validated against primary sources and current counsel guidance before reliance.