MUSE BLOSSOM

COMPLIANCE GUIDE · EU + US EDITION · AUGUST 2026

The EU AI Act Article 50 Compliance Guide

Machine-readable marking of AI-generated content: what the law requires, which technologies actually survive the real world, and a step-by-step checklist to get compliant. Now including the U.S. landscape: California SB 942, the NO FAKES Act, and the TAKE IT DOWN Act.

Aug 2, 2026
EU Art. 50 & California SB 942 apply

Dec 2, 2026
Grace period ends (existing systems)

€15M / 3%
Maximum fines for violations

Written by the team behind Korea's first officially C2PA-conformant product

museblossom.com

Article 50, at a glance

Article 50 of the EU AI Act sets transparency obligations for AI systems that interact with people or generate content. For generative AI, one clause matters most.

The core obligation (Art. 50(2))

Providers of AI systems that generate **synthetic audio, image, video, or text** must ensure outputs are **marked in a machine-readable format** and **detectable as artificially generated or manipulated**. Marking must be effective, interoperable, robust, and reliable — as far as technically feasible.

Who is covered

Providers — anyone developing a generative AI system and placing it on the EU market, or putting it into service in the EU, regardless of where the company is established. A Korean, US, or UK company serving EU users is in scope. **Deployers** carry separate disclosure duties for deep fakes and AI text published on matters of public interest (Art. 50(4)).

When it applies

DATE	WHAT HAPPENS
Aug 2, 2026	Article 50 obligations apply to generative AI systems on the EU market.
Dec 2, 2026	End of the grace period for systems placed on the market before Aug 2 — machine-readable marking must be in place.
Ongoing	An EU Code of Practice on marking and labelling of AI-generated content operationalizes the requirements; watermarking and provenance metadata are the reference techniques.

What non-compliance costs

Administrative fines

Transparency violations carry fines of up to **€15,000,000 or 3% of total worldwide annual turnover**, whichever is higher (Art. 99(4)).

The quieter costs

Platform distribution requirements, enterprise procurement questionnaires, and press scrutiny increasingly ask the same question: "how are your outputs marked?"

⚠ This guide is a practical engineering companion, not legal advice. Validate your specific obligations with counsel.

02 · SCOPE CHECK

Does Article 50(2) apply to you?

Three questions determine whether the marking obligation lands on your roadmap.

1 · Does your system generate synthetic audio, images, video, or text?

▼ yes

2 · Is it placed on the EU market, or are outputs used in the EU?

▼ yes

3 · Do any exceptions apply? (assistive/standard-editing functions, authorized law-enforcement use)

Exception applies

Document the analysis and keep it on file — scope can change with each model update.

No exception → you must mark

Machine-readable marking + detectability, by Aug 2 / Dec 2, 2026.

Edge cases teams get wrong

"We're just an API"

If you provide the generative system, you are the provider. Downstream apps building on your API inherit their own duties, but yours don't disappear.

"Our model only edits"

The assistive-editing exception is narrow: functions that don't substantially alter input data. Inpainting, face swaps, and style transfer usually don't qualify.

"We're not an EU company"

Irrelevant. The Act follows the market, not the company's registered address — the same extraterritorial logic as GDPR.

"Text is exempt, right?"

No — synthetic text is in scope for 50(2). A narrower deployer-side exception exists for human-reviewed editorial content, but that doesn't remove provider marking.

Marking technologies, compared

The regulation names the goal — effective, interoperable, robust, reliable — but not the method. In practice there are three families, and they fail differently.

Provenance metadata (C2PA)

Cryptographically signed "Content Credentials" attached to the file: who made it, with what tool, how it was edited. The open standard adopted by major platforms, browsers, and camera makers.

Invisible watermarking

A signal embedded in the pixels or audio itself, imperceptible to humans, readable by a detector. No visible change to the content; survives transformations that destroy metadata.

Visible labels

On-screen "AI-generated" badges and disclosures. Required in some deployer contexts (deep fakes), but trivially cropped — never sufficient alone for 50(2).

Survivability in the real world

EVENT IN THE CONTENT'S LIFE	C2PA METADATA	INVISIBLE WATERMARK	VISIBLE LABEL
Platform re-encodes upload	Usually stripped	Survives	Survives
User takes a screenshot	Gone	Survives	If in frame
Sent through a messenger app	Usually stripped	Survives	Survives
Cropped / resized	Gone if stripped	Survives (robust schemes)	Cropped out
Verifying rich provenance (who/how)	Full manifest	ID only	None
Interoperable, open ecosystem	Open standard	Vendor detector	Convention

The takeaway

Metadata gives you **interoperability and rich provenance**; watermarking gives you **survivability**. Each alone has a failure mode regulators and platforms already know about. Serious implementations ship both.

04 · REFERENCE ARCHITECTURE

The hybrid marking pipeline

What a compliant output path looks like when C2PA credentials and invisible watermarking work together.

1 · Embed

At generation time: sign a C2PA manifest (model, timestamp, edit chain) and embed an invisible watermark into the pixels / audio. No visible change, <100ms per asset in a well-built pipeline.

→

2 · Detect

Detection API finds the watermark even after compression, cropping, re-encoding, or screenshots — and flags AI-generated content at ingestion, moderation, or audit time.

→

3 · Verify

Anyone downstream validates the C2PA credential against the trust list; where metadata was stripped, the watermark re-links the file to its provenance record.

Design decisions that matter

- [ ] **Mark at the source, not post-hoc.**
  Embedding inside the generation pipeline covers every output path — API, app, batch — and can't be forgotten by a downstream team.
- [ ] **Choose a watermark benchmarked against your real distribution.**
  Ask vendors for detection rates after JPEG re-compression, resize, crop, and re-recording — not lab-condition numbers. False-positive rates matter as much as recall.
- [ ] **Prefer conformant C2PA over "compatible" C2PA.**
  The C2PA conformance program independently verifies that credentials interoperate with the ecosystem. Self-declared compatibility is exactly the claim you can't take to a regulator.
- [ ] **Don't forget audio.**
  Voice clones and AI music carry the same obligation. Inaudible watermarking exists and survives re-encoding — audio-only providers are the least prepared segment we see.
- [ ] **Log everything.**
  Marking records are your compliance evidence. Store manifest IDs, watermark payloads, and timestamps so any output can be traced back on request.

The 10-step compliance checklist

PHASE 1 · SCOPE & DECIDE — WEEK 1

- [ ] **1. Inventory every synthetic output surface.**
  Products, APIs, internal tools, marketing pipelines. Each modality (image / video / audio / text) gets its own row.
- [ ] **2. Run the scope check (page 3) per surface, with counsel.**
  Document conclusions — including the surfaces you decide are out of scope, and why.
- [ ] **3. Pick the marking architecture.**
  Hybrid (C2PA + invisible watermark) is the defensible default. Decide build vs. buy honestly: robust watermarking is years of R&D, not a sprint.

PHASE 2 · INTEGRATE — WEEKS 2–4

- [ ] **4. Wire marking into the generation pipeline.**
  One integration point at the source beats N integration points at the edges.
- [ ] **5. Validate robustness on your own content.**
  Re-compress, crop, screenshot, re-upload — then measure detection. Keep the benchmark report.
- [ ] **6. Validate C2PA output against the ecosystem.**
  Credentials must verify in standard tools, not just your vendor's demo.
- [ ] **7. Stand up detection for your own trust & safety loop.**
  Marking without detection is half the obligation — outputs must be *detectable*.

PHASE 3 · OPERATIONALIZE — ONGOING

- [ ] **8. Update user-facing disclosures.**
  Terms, API docs, and deployer guidance: tell downstream users the outputs are marked and how to verify.
- [ ] **9. Create the compliance evidence file.**
  Architecture description, benchmark results, conformance certificates, marking logs — the folder you open when a regulator or enterprise customer asks.
- [ ] **10. Re-test on every model release.**
  New model, new output distribution, new robustness numbers. Make marking validation part of the release checklist.

06 · BEYOND THE EU — THE UNITED STATES

The U.S.: same date, different laws

There is no single federal "Article 50" in the United States — instead, one state marking mandate that now mirrors the EU, plus federal takedown regimes.

California SB 942 — the AI Transparency Act

Generative AI systems **publicly accessible in California with over 1M monthly users** must embed **latent disclosures** in AI images, video, and audio (provider name, system + version, timestamp, unique identifier — permanent or extraordinarily difficult to remove), offer a **free public AI-detection tool** (web + API), and support an optional visible label. As amended by AB 853, it applies from **August 2, 2026 — deliberately aligned with EU Article 50**.

EU Article 50 vs. California SB 942

DIMENSION	EU AI ACT · ART. 50	CALIFORNIA · SB 942
Who is covered	Providers of generative AI on the EU market, any size	GenAI systems accessible in CA with >1M monthly users
Core duty	Machine-readable marking + detectability	Latent disclosure with named fields + free detection tool
Modalities	Audio, image, video, text	Audio, image, video (text excluded)
Applies from	Aug 2, 2026 (Dec 2 for pre-existing systems)	Aug 2, 2026
Penalties	Up to €15M or 3% of global turnover	$5,000 per violation — **each day counts separately**

The California timeline keeps going

Aug 2, 2026

Covered GenAI providers: latent disclosures + free detection tool live.

Jan 1, 2027

Large platforms must detect and surface C2PA-style provenance to users.

Jan 1, 2028

Capture devices sold in CA must offer provenance embedding at capture.

Why this matters: one hybrid integration — C2PA-conformant credentials + invisible watermark + detection API — satisfies the marking core of both regimes; SB 942's named disclosure fields map directly onto a C2PA manifest.

Likeness, deepfakes, and the takedown regime

Two federal instruments target synthetic likenesses rather than marking. Neither mandates watermarks — both make large-scale detection an operational necessity.

NO FAKES Act (pending — strong momentum)

Creates a federal, licensable property right in every individual's **voice and visual likeness** against unauthorized digital replicas. DMCA-style notice-and-takedown with a new counter-notice procedure; safe harbor for services that remove promptly. Reintroduced May 2026; advanced **unanimously** out of Senate Judiciary in June 2026 — its strongest position in four attempts.

TAKE IT DOWN Act (law since May 2025)

Requires covered platforms to remove nonconsensual intimate imagery — explicitly including AI-generated deepfakes — within **48 hours** of a valid request, with platform obligations live since May 2026. Enforced by the FTC.

What this means in practice

For platforms & hosts

Notice-and-takedown at scale is a detection problem: identifying voice clones and synthetic media on ingestion, blocking re-uploads of removed replicas, and distinguishing *authorized* replicas from infringing ones. Manual review does not survive the volume.

For rights holders & studios

A property right is only as strong as your ability to find infringements. Labels, studios, and talent agencies need continuous monitoring for unauthorized voice and likeness use — and, when they license authorized replicas, a way to mark and trace them.

One stack, both sides of the problem

Provenance credentials prove what is authentic and authorized. Invisible watermarks make licensed replicas traceable. Detection — of AI-generated music, cloned voices, and manipulated media — finds what shouldn't be there. The same three primitives that answer Article 50 and SB 942 also operationalize NO FAKES readiness.

State layer: Tennessee's ELVIS Act already protects voice against AI cloning, and dozens of state deepfake statutes are in force. U.S. legal exposure is a patchwork — but the technical answer converges.

07 · QUESTIONS WE HEAR WEEKLY

FAQ

We launched before August 2026 — are we off the hook?

No, you have a deferral, not an exemption. Systems placed on the market before Aug 2, 2026 must meet the machine-readable marking requirement by

December 2, 2026

. Four months disappears quickly once integration and robustness testing are on the critical path.

Is C2PA legally required?

The Act is technology-neutral — it requires machine-readable marking, not any named standard. In practice, C2PA is the interoperability layer the ecosystem is converging on and features prominently in Code of Practice discussions. Using a conformant implementation is currently the strongest interoperability evidence available.

Watermarks can be attacked. Does that make them pointless?

The obligation reads "as far as technically feasible" — it demands state-of-the-art robustness, not unbreakability. A benchmarked watermark that survives ordinary distribution (compression, cropping, screenshots) plus signed metadata is a defensible technical answer. "We did nothing because nothing is perfect" is not.

What about content we generated before the deadline?

Obligations attach to outputs generated by systems in scope going forward. But platforms and enterprise customers increasingly ask about back-catalog provenance too — batch marking of legacy libraries is a common second project.

Do open-source models change anything?

Releasing under an open license doesn't remove provider obligations for systems placed on the EU market. If you host an open model and serve EU users, the marking duty follows the service.

How long does integration actually take?

With an API-based vendor: typically 1–3 weeks to production, dominated by testing and rollout rather than code. Building in-house: budget quarters, not weeks — robustness engineering and false-positive tuning are the long poles.

08 · ABOUT

Why teams work with Muse Blossom

Certified, not self-declared.

In June 2026, Muse Blossom's Contents Defence became the first Korean product certified under the official C2PA conformance program — independently verified interoperability with the Content Credentials ecosystem. We pair those credentials with invisible watermarking that has spent five years being benchmarked against real-world piracy, not lab conditions.

Image Defence
Invisible watermarking + C2PA manifests for images, webtoons, and video. Leak tracing down to the account and channel.

Audio Defence
Inaudible watermarking for music and voice — ownership proof and synthetic-audio marking in one signal.

MusicScanner
Detection of AI-generated music at scale, for platforms and rights holders.

Integration
REST API and batch pipeline. Marking embedded at the generation source; detection as a service.

Founded 2021 · GS software quality certification · TIPS deep-tech R&D program · Offices in Changwon & Seoul, Korea · EU & overseas inquiries answered in English within 24h (CET-friendly).

See it on your own outputs.

A 15-minute demo: we mark and detect your sample content, live.

contact@museblossom.com