

Some AI tools now place a hidden marker on the content they generate, so that content can be identified as AI-generated. Companies including Anthropic and Google aligned this practice with the European Union (EU) AI Act, a transparency rule that took effect on August 2, 2026.

1. AI watermark basics

To understand what an AI watermark is, it is helpful to say what it is not. It is not a visible stamp, a footer, or a line of fine print. It is not the kind of watermark that becomes visible when you hold a paper up to the light. In fact, an AI watermark is not visible to the human eye.

There are two types of AI watermarks.

The **text watermark** is an invisible pattern stitched into the words as the AI tool writes them.

- When AI-generated content is copied, pasted, or forwarded, the mark is carried along with the words.
- It survives light human editing.
- It weakens as AI-generated text is heavily edited, paraphrased, translated, or integrated with other writing; sometimes it disappears completely. Other times it is reinforced, as AI continues to edit and regenerate text.

The **file label** is a small digital record attached to generated files, such as images, audio, or video. The label identifies the file as AI-generated.

What does an AI mark mean?

A mark means only that AI was involved in producing the words. It makes no claim of authorship or ownership of the ideas, judgment, or work. The mark identifies AI involvement, not a person.

Can anyone see the mark?

No. The watermark is unseen by the writer and by the reader. A watermarked document looks exactly like an unmarked document. Only detection software can see it.

Can I opt out or turn it off?

Where text marking is applied, there is no opt-out, because it happens at the AI model level.

Does the absence of a watermark mean AI was not used?

Not necessarily. The content could have been created by an AI tool that does not mark text, or the original mark could have been eliminated with editing.

Does the watermark indicate wrongdoing?

No. A mark means only that AI was involved in producing the words.

What can the watermark prove?

Less than most people assume. A mark is a transparency signal, not proof. It can show that AI was likely involved, but it cannot reliably identify someone who takes steps to remove it.

2. Purpose of the AI watermark

Public trust and legal compliance are two important factors in the push to establish AI transparency.

How does the mark support trust?

In a flood of synthetic content, people should be able to recognize AI-generated content. The AI mark is one step toward that, and it is aimed at misinformation, deepfakes, impersonation, fraud, and manipulated evidence.

How does the mark support legal compliance?

The EU rule requires AI-generated content to be identifiable. A hidden mark satisfies that requirement without altering the content, changing how it reads, or interfering with how it is used. Because the mark is machine-readable, an AI company can show that its output is identifiable, which is what the rule asks of it.

3. AI transparency in practice

Some companies chose to comply with the EU rule, in part because AI services are not confined by geographic borders. Rather than apply the marking only in Europe, companies that complied did so on a worldwide basis.

While the EU created the transparency rule, each AI company responded individually; this explains the variation in what content is marked and how it is marked.

For example:

- Claude (Anthropic) and Gemini (Google) mark the text they generate.
- ChatGPT (OpenAI) does not mark text. It does mark generated images and audio, though those marks are not always hidden.

Many AI products do not build their own models. They route to models built by other companies and inherit whatever those models do. Copilot (Microsoft), for example, routes to models from more than one company, so what is marked depends on which model is behind a given response.

Nearly 190 companies and organizations have signed the EU's voluntary Code of Practice on marking AI-generated content, but commitment has moved faster than deployment.

4. Future of watermark use and detection

Where does detection stand today?

Detection of AI watermarks is far less available than headlines suggest. Official detectors from the AI companies are still in development and not broadly available; third-party checkers are unreliable and unproven.

What comes next?

Expect the use of AI watermarks to increase. More tools and more content types are likely to be included over time, though the pace and the details are not settled. Expect detection tools to become more easily available as well. Though no release dates have been announced, AI companies are building their own detectors.

5. Impact on your work

Will my work have a watermark if I use AI but edit heavily as I do so?

Most likely, yes. Heavy editing, paraphrasing, or rewriting by you weakens the mark and can eventually remove it; however, if your process includes asking AI to revise or regenerate along the way, each iteration carries a new mark. The mark is more likely to remain.

Does a mark mean I have to disclose that I used AI?

Not by itself. Marking and disclosure are two different things. Marking happens automatically when AI generates text. Whether you need to disclose depends on your situation and on what a client, employer, court, publisher, or licensing body expects of you. Your content may well carry a mark and carry no obligation to disclose anything.

Should I disclose my AI use voluntarily?

That is your call, and there is no rule requiring it in most settings. Some people add a short line noting that AI assisted the work, which signals confidence rather than apology. Others disclose only where it is expected. Either approach is defensible; what matters is that you are consistent.

Can I know before publishing whether my content carries an AI mark?

Not yet. As noted above, the tools that would let you check your own work are still in development.

Sources and further reading

- Claude Help Center, "How Claude marks AI-generated content." support.claude.com/en/articles/16266773
- Microsoft Learn, watermarking for AI-generated content in Copilot.
- EU Artificial Intelligence Act, Article 50 transparency obligations. artificialintelligenceact.eu/transparency-rules-article-50/

About this resource. AI content watermarking took effect in August 2026, and shared understanding of it is still developing. This document reflects information available at the time it was written and may change as tools, detection methods, and rules evolve. It is educational and is not legal or compliance advice. Authored by AI Efficiency Labs in conjunction with the drafting and editing capabilities of Claude, an AI assistant by Anthropic.