PLAIN-ENGLISH GUIDE

OCTOBER 2026 · CHATGPT · CODEX · TEXTGRAIN

AI WATERMARKS WHAT THEY MEAN FOR YOU

ChatGPT is starting to hide an invisible watermark in the words it writes. This guide explains what OpenAI's textGrain is, who gets it, how well it works, what it can and can't reveal, and what it means if you write with AI at school, at work or for clients.

01
What changed
Who gets the watermark, and when.

02
How it works
A pattern in word choices, not hidden characters.

03
What it means
For students, freelancers, teams and developers.

01 WHAT CHANGED

ANNOUNCED OCT 5, 2026 · EU AI ACT

OpenAI announced **textGrain**, an invisible watermark for AI-written text, to help label AI content under the EU's AI Act. It is not a global default.

WHO GETS IT

WHERE	WHAT HAPPENS	WHEN
ChatGPT in the EU	Eligible text output gets the watermark, on **all plans**	Rolling out "over the coming weeks" from Oct 5
Codex in the EU	Eligible output gets the watermark	Same rollout
API, worldwide	Developers can switch it on for select models. **Off by default.**	Available from Oct 5
Everyone else	No watermark by default	No date announced

WHO CAN CHECK FOR IT

- OpenAI's detector is **not public**.
- At first only **approved researchers and expert organizations** can apply for access.

DOES IT CHANGE THE ANSWERS?

- OpenAI says it saw no meaningful change in quality with the watermark on.
- You won't see anything different on screen.

02 HOW IT WORKS

A STATISTICAL SIGNAL IN WORD CHOICE

Every time a model writes, it picks each next word from several good options. textGrain gently nudges those picks so that, over many words, a hidden statistical pattern builds up. A reader can't see it. A detector that knows what to look for can.

THINK OF IT LIKE THIS

- "The results were **clear**" and "The results were **obvious**" both read fine.
- The watermark leans the choice one way. One word proves nothing; hundreds of leaned choices add up to a pattern.
- That's why longer text is easier to detect than a single sentence.

WHAT IT IS NOT

- Not hidden characters, invisible spaces or odd punctuation: it lives in the words themselves.
- Not visible when you read or copy the text.
- Not a label on your account: it says nothing about who wrote the prompt.

WHAT A DETECTED WATERMARK DOES NOT TELL ANYONE

- **Who you are**, your account, your prompt or your conversation.
- **How much a human wrote or edited** in the final piece.
- Who owns the text, who is responsible for it, or whether it is accurate.

03 HOW RELIABLE?

OPENAI'S OWN TESTS · ~400-TOKEN PASSAGES

OpenAI published detection results with the announcement. On passages of about 400 tokens (a few paragraphs), the detector caught the watermark about 92% of the time. Editing the wording weakens it fast.

~92%
Unedited
a few paragraphs straight from the model

66%
10% reworded
one word in ten swapped for a synonym

17%
25% reworded
a quarter of the words swapped

?
Hard cases
short text, maths answers, translations

THE TWO BIG RULES

1. **No watermark doesn't prove a human wrote it.** The text may be too short, edited, translated, or made by another company's AI.
2. **Mistakes happen both ways.** OpenAI says false positives and false negatives both occur, especially on short texts and subjects like maths.

! Because the detector isn't public, any website claiming to "check ChatGPT's watermark" for you is **not OpenAI's detector**. Treat AI-text checkers as rough guesses, never as proof.

04 WHAT IT MEANS

STUDENTS · CREATORS · TEAMS · DEVELOPERS

STUDENTS

Know your rules

The watermark doesn't change your school's AI policy: it can only make AI text easier to spot later. If AI help is allowed, say how you used it. Keep your notes and drafts.

FREELANCERS

Agree up front

Tell clients when AI drafts are part of your process and where your own work comes in. Clear terms beat surprises later.

TEAMS

Write it down

Add a short AI-use line to your content policy: which tools, which jobs, who reviews, and how AI help is disclosed.

DEVELOPERS

It's your switch

In the API (worldwide, select models) the watermark is `off` by default. Decide on purpose whether your product turns it on.

COPY-PASTE: AN AI-USE NOTE FOR YOUR WORK

```
AI use: I used [tool] to [brainstorm / outline / draft / edit]
[which parts]. I checked the facts, rewrote [which parts] in my
own words and I'm responsible for the final version.
```

COPY-PASTE: ASK CHATGPT ABOUT YOUR SITUATION

```
I [study at / work for / freelance for] [where]. Summarise what
OpenAI's text watermark (textGrain) does and doesn't reveal, then
help me draft a short, honest note on how I use AI for [task].
```

DO THIS WEEK

CHECKLIST

1. **Check if it affects you**
   In the EU and using ChatGPT or Codex? Eligible text gets the watermark as it rolls out. Elsewhere it's off by default.

2. **Read the rules you're under**
   School, employer or client AI policies matter more than any watermark. Ask if there isn't one.

3. **Be open about AI help**
   A one-line AI-use note (page 5) is the simplest way to stay on the right side of any policy.

4. **Keep your drafts**
   Notes, outlines and version history show your own work far better than any detector.

5. **Don't judge others by a checker**
   No watermark proves nothing, and public "AI detectors" make mistakes. Never accuse someone on a score alone.

6. **Watch for other AIs**
   Anthropic announced its own text watermark for Claude in August. Expect more of this as AI labelling rules arrive.

STILL UNKNOWN

- When (or whether) the watermark comes to ChatGPT outside the EU by default.
- When the detector opens up beyond approved researchers.

CUT THROUGH THE AI NOISE

textGrain was announced on October 5, 2026 and is still rolling out. Details can change: check OpenAI's post for the latest.

news.aistacksignal.com/watermark

SOURCES (CHECKED OCT 27, 2026)

OpenAI announcement: openai.com/index/eu-text-provenance
textGrain technical report: cdn.openai.com (PDF)
TechCrunch: OpenAI will start watermarking ChatGPT's text in the EU
BleepingComputer: OpenAI is adding invisible watermarks to ChatGPT and Codex text in the EU
9to5Mac: OpenAI details new text watermarking system

AI Stack Signal is not affiliated with OpenAI or Anthropic. Figures are OpenAI's own test results as reported on the date above. This guide is general information, not legal advice; check the rules that apply to you.