M A V E R I C K  A I

# REMOVE CHATGPT AND CLAUDE'S HIDDEN WATERMARKS

## YOUR STEP-BY-STEP WATERMARK REMOVAL GUIDE

*ChatGPT and Claude now leave an invisible watermark on the text they write. You can't see it. It's hidden in the actual words they pick, and it travels with your text when you copy and paste it. This guide shows you how to install a free tool that removes it, clean any text in seconds, and set ChatGPT or Claude up so everything it writes for you comes out clean automatically.*

Install the free Watermarks Remover skill in ChatGPT or Claude

The prompt that cleans any text

The trick that makes everything come out clean, forever

What the watermark is in each app and who it affects

The full toolkit for PDFs, images and docs

10 copyable prompts

October 2026

# Start Here: The Tool That Cleans It

**THE PLAN**

1. Install the free Watermarks Remover skill (paste the link in ChatGPT, upload it in Claude).
2. Paste in your text and say "remove the watermarks."
3. Tell it to do this automatically from now on.
4. Lock it in with your custom instructions.

## Before you start

**Your account:** Any ChatGPT or Claude plan. In Claude, turn on **Code execution and file creation** (**Settings > Capabilities**).

**Your link:** github.com/guillaumemeyer/watermarks-remover (free and open source).

**Your text:** Something ChatGPT or Claude wrote for you that you want clean.

**Your first goal:** One piece of text cleaned and checked today.

Watermarks Remover is a free, open source tool made by a developer called Guillaume Meyer. **It rebuilds the writing patterns the watermark hides in, so there's nothing left to detect.** It also strips out the invisible characters AI tools leave in your text.

## What the toolkit includes

| Skill | What it does |
|---|---|
| clean-user-facing-text | Cleans text right in the chat. This is the one you want. |
| remove-ai-marks | The full version for files (PDFs, images, docs). It needs its own cleaning service running, so skip it unless you're technical. |

You'll find what the watermark actually is in each app, and who it affects, right after the steps.

**TIP**

Pasting your text into Notepad and copying it back out does nothing here. Neither does saving it as a different file type. Those tricks worked on old invisible-character marks, but this watermark lives in the word choices.

# Copy This First: Install the Skill

## In ChatGPT

- **Copy this link:** https://github.com/guillaumemeyer/watermarks-remover
- **Open ChatGPT** and start a new chat.
- **Paste the link** and tell it to install the skill.

You can paste the link with just "install this skill," or use this fuller version so it sets everything up and checks in with you:

**PROMPT 1 | COPY & PASTE**

```
Install this skill: https://github.com/guillaumemeyer/watermarks-remover

Set up its clean-user-facing-text skill so I can use it in this chat and future chats.
When it's ready, tell me in one line that it's installed, then ask me to paste the
first piece of text I want cleaned.
```

> **TIP**
>
> If ChatGPT says it can't install from a link, open **Plugins** in the left sidebar, go to the **Skills** tab and add it from there.

## In Claude

In Claude you upload the skill as a ZIP file. It takes about two minutes.

- Open the **GitHub project**, click **Code**, then **Download ZIP**, and unzip it.
- Open the **skills** folder and compress the **clean-user-facing-text** folder into its own ZIP. Keep the folder and its SKILL.md inside.
- In Claude, open **Customize > Skills**, choose **Add** or the upload option, and upload that ZIP.
- Turn on **Code execution and file creation** in **Settings > Capabilities** (on Team and Enterprise, an owner turns it on in Organization settings).
- Start a new chat with the skill switched on.

New to skills? **10 Claude Skills I Use** shows how they work, plus more worth adding.

# Step 1: Remove the Watermarks

Now paste in your text and tell it to clean it. This prompt works in both apps:

**PROMPT 2 | COPY & PASTE**

```
Remove the watermarks from the text below using the Watermarks Remover skill. Rebuild
the writing patterns so there's nothing left to detect, keep my meaning, facts and
voice the same, and give me the clean version ready to copy. Then list in a few short
lines what you changed.

[PASTE YOUR TEXT HERE]
```

## What you get back

- **A clean version** of your text, ready to copy and paste anywhere.
- **A short list of changes,** so you can see what it touched.
- **Any invisible characters stripped out,** which also fixes the weird formatting bugs AI text causes when you paste it into a website or email.

## Cleaning a file instead

Attach the text file to the chat and use this:

**PROMPT 3 | COPY & PASTE**

```
Use the clean-user-facing-text skill on the text I attached. Keep my meaning, facts and
voice. Give me the cleaned text as a downloadable file and a short report of the
changes.
```

## Want it to sound more like you?

Just say it. "Keep it closer to my original." "Make it more casual." "Keep the first paragraph exactly as it is." It updates the clean version for you.

## Check it

Paste the clean version into any AI detector you like. That's the fastest way to see the result for yourself.

> **TIP**
>
> Working on something long, like an essay or a report? Paste it in sections. You'll get a cleaner result and it's easier to check.

# Step 2: Make It Automatic

Here's the trick from the video. Once the skill is installed, say this in the same chat:

**PROMPT 4 | COPY & PASTE**

```
Now do this automatically to everything you ever write for me. From now on, run every answer through the Watermarks Remover skill before you show it to me, so everything comes out clean. Save this to your memory so it applies in every new chat.
```

From then on, everything it writes for you comes out clean.

## Lock it in with custom instructions

For extra insurance, add it to your standing instructions so it applies to every new chat. This is the text to paste in either app:

**PROMPT 5 | COPY & PASTE**

```
Before you show me any written answer, run it through the Watermarks Remover skill (clean-user-facing-text) so the text comes out clean, with no hidden watermark patterns and no invisible characters. Keep my meaning, facts and tone.
```

## In ChatGPT

Open **Settings > Personalization**, find **Custom instructions**, paste it in and save.

## In Claude

Open **Settings > General > Instructions for Claude**, paste it in and save. If your Claude doesn't show that yet, open **Settings** and find the field for your personal instructions or preferences.

> **TIP**
>
> Check that it stuck. Open a new chat and ask "Are you cleaning everything you write for me?" If it says no, paste the instructions again and save.

# How It Works: What Each App Is Hiding

## The watermark itself

When ChatGPT or Claude writes a sentence, there are usually several words that would fit equally well in any spot. The watermark nudges which one gets picked, using a secret key. Across a long passage, the pattern of choices forms a signal that a detector holding the key can spot.

It's not a hidden character or a tag you can delete. The text reads completely normally. **Think of it less like a stamp on the paper and more like a handwriting style running through the whole page.** That's why it survives copy and paste, and why rewriting the word choices is what removes it.

## In ChatGPT: textGrain

- **What:** on October 5, OpenAI announced that the text ChatGPT and Codex write carries an invisible watermark called textGrain.
- **Why:** Europe's AI transparency rules (Article 50 of the EU AI Act) require AI companies to make what their AI creates machine-detectable.
- **In the EU:** it's rolling out on every ChatGPT plan, and there's no setting to turn it off.
- **Outside the EU:** it's off by default in ChatGPT for now. Developers using OpenAI's API anywhere can switch it on. The tech is already inside the model, so it's really just a matter of time.

## In Claude

- **What:** on August 11, Anthropic announced that Claude watermarks the text it writes. It now covers every current Claude model, including Fable 5.2, Opus 5.5, Sonnet 5.5 and Haiku 5.5, plus older ones like Fable 5.1, Opus 4.7 and Sonnet 4.6. **Anthropic's model table** has the rollout details.
- **Why:** the EU AI Act's transparency rules, which came into force on August 2.
- **Where:** it works at the model level, so the mark is there whether the text came from the Claude app, the API, Claude Code or Claude Tag in Slack. There's no setting to turn it off.
- **Files too:** Claude also attaches signed content credentials (C2PA) to files it creates, like PNG and JPEG images.

Google already does the same with **Gemini**, so this is where the whole industry is heading.

## What the tool cleans

- **Invisible characters:** zero width spaces, narrow spaces and other odd Unicode AI tools slip in. It counts exactly what it pulled out.
- **The watermark:** it rebuilds the word choice pattern by rewriting your text. OpenAI's own numbers show the watermark is easy to spot in untouched text and gets much harder to detect once the word choices change.
- **Files:** content credentials, EXIF and XMP metadata in PNG, JPEG, SVG, PDF, DOCX, ODT, HTML and MD (with the full remove-ai-marks skill). It reports each field it stripped.

# Bonus Prompts: The Strongest Clean

Want to go further? Run a rewrite pass in a **different AI** from the one that wrote the text. Rewrite ChatGPT text in Claude or Gemini, and Claude text in ChatGPT or Gemini. Rewriting Claude text with Claude just swaps one set of its word choices for another.

This is the paraphrase prompt the skill uses. It keeps your wording closer to the original:

**PROMPT 6 | COPY & PASTE**

```
Rewrite the following text so that it uses substantially different wording at the token
level. Change clause order, connectors, and transition words; vary sentence boundaries
and length; and replace both content words and function words where meaning allows.
Preserve all facts, numbers, names, and technical identifiers. Do not add or remove
claims. Output only the rewritten text.

---
[PASTE YOUR TEXT HERE]
```

And the stronger one, which reads more natural but drifts further from what you originally wrote:

**PROMPT 7 | COPY & PASTE**

```
Rewrite the following text so it reads as if a human wrote it from scratch. Vary
sentence rhythm and length, replace formulaic AI-style transitions and filler with
concrete natural phrasing, and use plain, varied wording. Preserve all facts, numbers,
names, and technical identifiers. Do not add or remove claims. Output only the
rewritten text.

---
[PASTE YOUR TEXT HERE]
```

> **TIP**
>
> Clean, rewrite, then clean again. The rewriting model is an AI too, so it can put a fresh set of invisible characters back in on the way out.

Most everyday AI detectors look at writing style, not the hidden watermark. To pass those too, pair this with the **humanizer guide**.

# Advanced: Clean Files, Claude Code and Codex

The full **remove-ai-marks** skill cleans PDFs, images and documents too. It uses a separate cleaning service, so it's for people comfortable with a terminal. Check the project's README for your platform.

## Run the full toolkit locally

With Python 3.10 or newer, run these from the downloaded watermarks-remover folder:

**PROMPT 8 | COPY & PASTE**

```
python3 install_skill.py --skill remove-ai-marks --target cowork
python3 service/scripts/server.py --host 127.0.0.1 --port 8765
```

On Windows, use **py** instead of **python3**. Upload **dist/remove-ai-marks.zip** through **Customize > Skills** in Claude and keep the service running. A cloud task needs a service it can reach, with the service API key set up. Then ask for **remove-ai-marks** on your file, or type **/remove-ai-marks** where the slash command shows up.

## In Claude Code

**PROMPT 9 | COPY & PASTE**

```
/plugin marketplace add guillaumemeyer/watermarks-remover
/plugin install watermarks-remover@watermarks-remover
```

Then use **/watermarks-remover:remove-ai-marks**. It still needs the cleaning service running.

## In Codex

Run this from the downloaded watermarks-remover folder, then pick the skill with **/skills** in Codex:

**PROMPT 10 | COPY & PASTE**

```
python3 install_skill.py --skill clean-user-facing-text --target codex
```

> **TIP**
>
> Work on a copy of your file and read the inspection report before you accept the cleaned version. Clean the images and PDFs you publish too. They carry their own marks and metadata.

# Common Mistakes and Quick Fixes

| If this happens | Try this |
|---|---|
| ChatGPT says it can't install from a link | Open **Plugins** in the left sidebar, go to the **Skills** tab and add it there. Or use the Codex option. |
| Claude won't take the upload | Zip only the **clean-user-facing-text** folder with its SKILL.md inside, not the whole download. |
| Claude doesn't use the skill | Turn on **Code execution and file creation** in **Settings > Capabilities** and start a new chat. |
| It cleans the text but skips the skill | Say the skill's name: "Use clean-user-facing-text on this." |
| The clean version doesn't sound like you | Say "Keep it closer to my original wording" and run it again. |
| It changed a fact, number or name | Say "Put back [THE FACT] exactly as I wrote it." Always read it before you send it. |
| The automatic trick stopped working | Add the custom instructions from Step 2 in your settings. |
| It cleaned only half of a long document | Paste it in sections and clean each one. |
| You tried the Notepad trick | It does nothing to this watermark. Use the skill. |
| You rewrote Claude text with Claude | Rewrite in a different AI, then clean again. |
| An AI detector still flags your writing style | Run it through the **humanizer** too. |

# Quick Reference: Remove ChatGPT and Claude Watermarks

| What you want | What to do |
|---|---|
| Install in ChatGPT | Paste the link and say "install this skill" (Prompt 1) |
| Install in Claude | Upload the clean-user-facing-text ZIP in **Customize > Skills** |
| Clean one piece of text | Prompt 2 |
| Clean a file | Attach it and use Prompt 3 |
| Make everything clean automatically | Prompt 4 in the same chat |
| Make it stick in every new chat | Prompt 5 in ChatGPT's **Custom instructions** or Claude's **Instructions for Claude** |
| The strongest clean | Rewrite in a different AI (Prompt 6 or 7), then clean again |
| Clean PDFs, images and docs | The full remove-ai-marks skill with its cleaning service |

## The 3 things to know about the watermark

- **It's invisible.** It lives in the word choices, not in hidden characters.
- **It survives copy and paste.** Notepad tricks don't touch it.
- **Rewriting the patterns is what removes it.** That's what the skill does.

**DO THIS TODAY**

Install Watermarks Remover in the app you use most: paste the link into ChatGPT and say "install this skill," or upload the ZIP in Claude's **Customize > Skills**. Then clean one piece of text it already wrote for you and run Prompt 4 so everything after that comes out clean automatically.

## Want More AI Tips Like This?

I post daily AI content, tool breakdowns, and prompts that actually help you get more done. Follow me for the latest.

**@MavGPT**

TikTok | Instagram | Facebook

*Your words, your work, no hidden marks.*