Free · no sign-up

The Claim Check

One prompt and a five-minute routine for checking any AI answer, article or draft before you rely on it.

AI tools write fluently, which makes their mistakes easy to miss. A wrong date or an overstated rule sits comfortably inside a good paragraph. In one study by Columbia University's Tow Center, eight AI search tools were asked to identify the source of news excerpts. They got it wrong more than 60% of the time, usually without any sign of doubt (Nieman Lab, March 2025).

The Claim Check breaks a piece of text into single claims and checks each one, so a mistake has nowhere to hide. Then it tells you which few claims to check yourself.

When to use it

- Before you forward, publish or act on an AI answer.
- On a draft written with AI help, before it goes out under your name.
- On an article or report someone sends you that seems too neat.

It's built for factual text: dates, figures, rules, names, quotes. It won't tell you whether an opinion is right.

How to use it

1. **Open a new chat.** Don't use the chat that wrote the text, because a model checking its own work tends to defend it. A different AI tool is even better.
2. **Copy the prompt below** and paste the text you want checked where it says `[[PASTE THE TEXT]]`.
3. **Read the summary first.** It lists the claims to fix and the high-stakes claims to check yourself.
4. **Do the five-minute check** (below) on those high-stakes claims.

If your AI tool can search the web, it will try to open the original source for each claim. If it can't, the prompt makes it say so and label every verdict "FROM MEMORY - NOT CHECKED". That label tells you what still needs checking.

The prompt

```
Check the factual claims in the text below. You are a careful fact-checker, not the
author. Do not rewrite the text.

TEXT TO CHECK:
[[PASTE THE TEXT]]

STEP 1 - LIST THE CLAIMS
Split the text into separate factual claims (one fact each). Ignore opinions and
advice.

STEP 2 - CHECK EACH CLAIM
If you can search the web: for each claim, find the original source (the organisation
that produced the figure or made the rule, not an article about it), open it, and
compare the wording.
If you cannot search the web: say so at the top, do NOT cite any URLs, and mark every
verdict "FROM MEMORY - NOT CHECKED".

Return one table:

| # | Claim | Verdict | Problem | Corrected wording | Source (title, URL, short exact
quote) |

Verdict is one of:
  CORRECT - the source says this.
  NEEDS QUALIFIER - true only with a limit the text leaves out (a date, a place, "up
to", "in one study").
  WRONG - the source says something different.
  OUTDATED - was true, but has changed.
  CAN'T VERIFY - you could not find or open a reliable source.

STEP 3 - SUMMARY
- The claims I must fix before using this text.
- The claims I should check myself because they are high-stakes (numbers, dates,
quotes, legal, medical or money matters).
- Anything important the text leaves out that a reader would need.

Rules: never invent a source, never cite a page you did not open, and prefer "CAN'T
VERIFY" to a guess.
```

What you get

A table with one row per claim:

Verdict	Means	What to do
CORRECT	The source says this	Keep it. Spot-check it if it's high-stakes
NEEDS QUALIFIER	True only with a limit the text leaves out	Add the qualifier ("in the UK", "up to", "in one study")
WRONG	The source says something different	Fix it or cut it
OUTDATED	Was true, has changed	Update it, with the date
CAN'T VERIFY	No reliable source was found or opened	Treat it as unproven until you've checked it yourself

Then a summary: what to fix, what to check yourself, and what the text leaves out.

Our test: we planted five errors

We wrote a 117-word paragraph about the EU AI Act containing **five deliberate errors**: two wrong dates, a wrong maximum fine, a false claim that the Act only applies to EU companies, and an overstated claim that it bans all facial recognition in public. It also contained three correct statements. We ran the Claim Check on Claude twice on 10 November 2026: once with web search allowed and once without.

Both runs flagged all five errors and corrected them, from the model's memory: neither run opened a single source (see below). They also left the correct dates alone, and added qualifiers we hadn't planted, such as the later deadline for general-purpose AI models already on the market.

The more useful result was the failure. In the web-enabled run, the AI could search but every page it tried to open was blocked by our test environment's network rules. It didn't bluff. It said at the top that it couldn't open any page, marked every verdict "CAN'T VERIFY", gave its best judgement from memory, and listed the official pages to check. The no-web run did the same, labelling everything "FROM MEMORY - NOT CHECKED". That's the behaviour the prompt is designed to force: **you always know what has and hasn't been checked.**

What we haven't tested: other AI tools (ChatGPT, Gemini, Copilot), other subjects, or longer texts. One test on one model, with no sources opened, isn't proof, and 5 of 5 here is not an accuracy rate. Treat it as a reason to try the prompt yourself, not as a guarantee. The full transcripts are published with the toolkit.

The five-minute check (no AI)

For every claim the summary calls high-stakes (numbers, dates, quotes, legal, medical or money matters):

1. **Open the original source.** That's the organisation that made the rule or produced the figure, not an article about it. If the AI gave a URL, open it. If it couldn't, search the organisation's own site.
2. **Find the exact words.** Press Ctrl+F (Cmd+F on a Mac) and search for a distinctive word from the claim.
3. **Compare carefully.** Look for missing qualifiers: "up to", "in England", "from April", "for firms that took part".
4. **Note the date you checked.** Rules, prices and AI products change within weeks.

If you can't find it in five minutes, the claim is unproven. Leave it out or say it's unconfirmed.

Privacy

Don't paste anything you wouldn't be comfortable showing the AI company: personal details about other people, confidential client material, unpublished financial figures. Replace names with roles ("Client A"). Your organisation's AI policy comes first.

Where this comes from

The Claim Check combines two steps from **Workflow 1: Research and fact-check** in our *AI Workflow Toolkit* (working title), which is in development. The full workflow goes further: it frames the question before any research, researches from a *source pack* of documents you choose so the AI can't drift, and runs the challenge in a separate chat. You check the sources yourself before writing a one-page brief with a confidence rating. The toolkit has two more workflows built the same way, for turning ideas into projects and for drafting and reviewing professional writing.

The toolkit isn't for sale yet. If you'd like to hear when it's ready, or you tried the Claim Check and want to tell us how it went, the Products page on our website is where it will appear first.

The Claim Check v0.1 · Understanding AI HQ · November 2026. Free to use and share unchanged, with credit. Tested on one AI model only (Claude); see above.

Understanding AI HQ · Free sample · from the AI Workflow Toolkit · 10 November 2026. Tested on one AI model only. See the testing notes for what was and wasn't tested.