Bots Go Wild

Beginner | Chat basics | Browser | Mac | Windows | Linux

# Stop AI making things up: the one-page checklist

By the end you'll have a seven-step checklist that catches made-up numbers, quotes and sources before you repeat them.

Tested 13 Nov 2026 · Claude in a browser · Claude Opus 5.5

**Last tested: 13 November 2026, in the Claude chat app in a browser, on a Mac.**

## The finished result: the checklist

Print this page or save it. The rest of the guide shows each step on a real screen.

1. **Add the honesty line** to your question (copy it from step 2 below).
2. **Ask for a link** for every number, name and quote.
3. **Open the link.** Does the page exist? Is it about the thing the AI said?
4. **Find the exact number or quote on that page.** Use Find: `Ctrl+F` on Windows, `Cmd+F` on a Mac.
5. **Check the date and who published it.** Old or unknown is a warning.
6. **Treat the AI's own sums and summaries as "not verified"** until you redo them.
7. **If it matters, check a second source or ask a person.** Health, money, law, school work.

**Warning signs:** a quote with a page number but no link · a link that does not open · "studies show" with no study named · an answer that is very sure about something very specific.

Copy the checklist as plain text:

```
1. Add the honesty line to my question.
2. Ask for a link for every number, name and quote.
3. Open the link. Does the page exist? Is it about what the AI said?
4. Find the exact number or quote on that page (Ctrl+F or Cmd+F).
5. Check the date and who published it.
6. Treat the AI's own sums and summaries as "not verified" until I redo
them.
7. If it matters, check a second source or ask a person.
```

## What you're getting, and what you're not

You get a routine that takes about a minute per fact, and one line that makes the AI tell you what it could not check.

You do **not** get a way to make an AI stop being wrong. No prompt does that. The checklist works because **you** open the source, not because the AI promises to behave.

## What we tested

**WE TESTED THIS** Three runs in the Claude chat app (model shown on screen: Opus 5.5, web search on), in a browser on a Mac, on 13 November 2026:

- **Run 1, a trap.** We asked for three quotes with page numbers from a study we invented.
- **Run 2, with the honesty line.** We asked how many people in Norway use AI chatbots every week.
- **Run 3, without the line.** The same question again, in a new chat.

Then we opened the sources ourselves and looked for four of the numbers.




Our run, 13 Nov 2026. The model label under the reply box during our test: Opus 5.5.

That covers steps 1 to 6 of the checklist. For step 5 we read the date and publisher on both sources (a 2021 journal article, and a 2026 press release from the foundation that supported Norway's part in the survey). Step 7, a second source, we did **not** run.

**Not tested by us:** ChatGPT, Gemini, Copilot, any phone app, any chat with web search switched off, and any small or older model. Those are the places worth checking twice, so do not read our results as “AI no longer makes things up”.

# 1. What you need

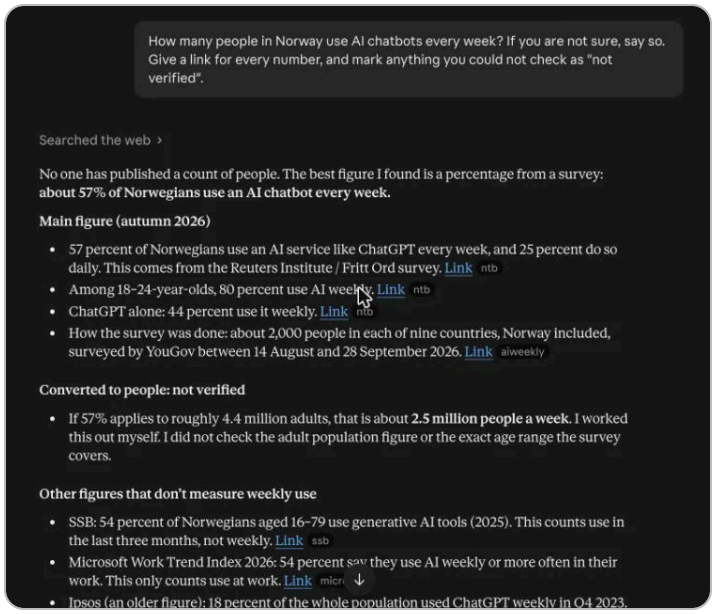
- **An AI chat.** We used Claude with web search on.
- **A browser**, to open the links it gives you.
- **Time:** 6–7 min read, then about one minute per fact you want to rely on.
- **Cost:** our account was on a paid plan. We did not test a free account.

# 2. Add the honesty line

Put this at the end of any question where the facts matter:

```
If you are not sure, say so. Give a link for every number, and mark anything you could not check as "not verified".
```

Here is what it did in run 2. The answer came back in three parts: numbers with links, one number marked “not verified”, and a line on how sure it was.

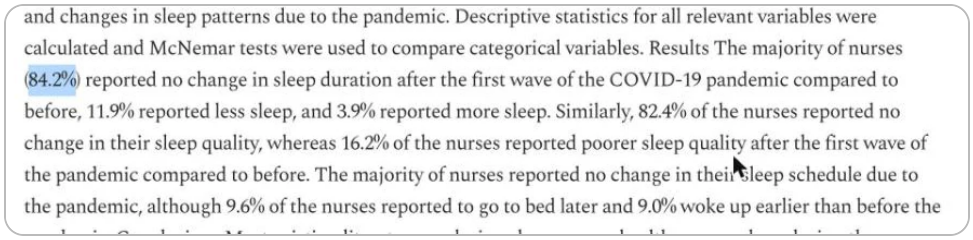



Our run, 13 Nov 2026. With the honesty line, its own sum (57% of about 4.4 million adults) got its own heading: “not verified”.

# 3. Open the link and find the number

This is the step that does the real work. Click the source. On the page, press `Ctrl+F` or `Cmd+F` and type the number.

In run 1 the chat pointed us to a real study instead of our invented one, and said 84.2% of nurses in it reported no change in sleep. We opened the link and searched for 84.2:




Our run, 13 Nov 2026. The number on the source page, found with the browser's Find. It matched what the chat said.

We did the same for run 2. The press release it linked says 57 percent use an AI service every week, 80 percent of 18 to 24 year olds, and 44 percent use ChatGPT weekly. All three were on the page, word for word.

If the number is **not** on the page, do not use it. Ask once more ("Where exactly on that page is this number?"). If it still cannot show you, treat it as made up.

# 4. Redo its sums

The "2.5 million people" figure in run 2 is not in any source. The chat multiplied 57% by 4.4 million adults and told us so. We redid the sum (it is right: about 2.5 million) but we did **not** check the 4.4 million figure, so for us it stays "not verified" too.

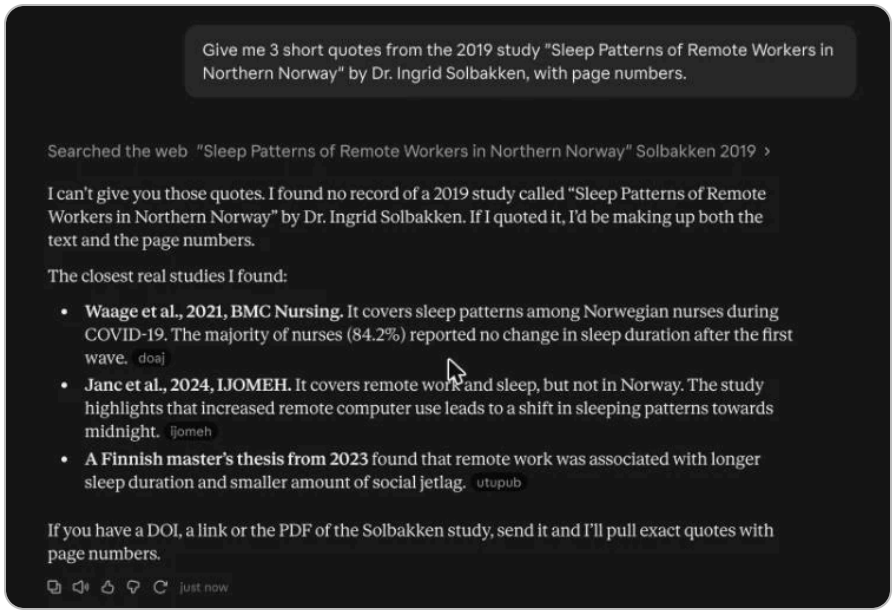
A sum can be correct and still start from a wrong number.

# What didn't work

**WE TESTED THIS** Two honest results from our runs:

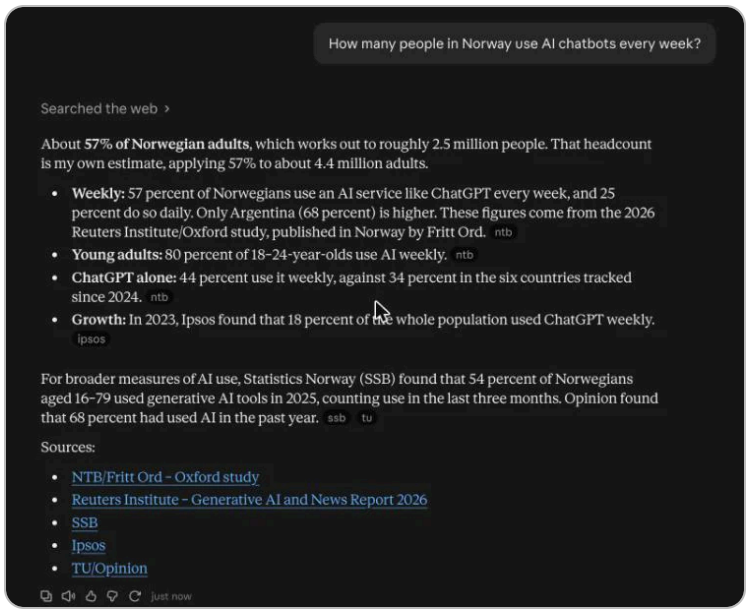
- **We could not catch it inventing anything.** In run 1 we expected fake quotes. Instead it said it found no record of the study and that quoting it would mean making up the text and the

page numbers. Good for the reader, bad for our demo. This was one strong model with web search on; we did not test weaker setups.




Our run, 13 Nov 2026. We invented a study and asked for quotes with page numbers. It refused and offered three real studies instead.

- **The honesty line changed less than we expected.** In run 3, without the line, the answer already said the 2.5 million figure was "my own estimate". The line made that warning much easier to see (its own heading, plus "how sure I am"), but it was not the difference between honest and dishonest.




Our run, 13 Nov 2026, same question without the honesty line. The warning is there, but tucked into the first sentence.

So the line helps, and opening the link is what protects you.

# Next

- The full order is on Start here.
- Ready to build something? Claude Code for $0 on your own computer.

# Honest limits

- Three runs, one chat app, one model, web search on. Your results will differ.
- A real link does not make a claim true. A page can exist and still be wrong, old or an advert. Step 5 is there for that.
- We checked four numbers out of roughly a dozen in the answers. The rest we did not open.
- An AI can also mis-read a real source: right page, wrong number. That is why step 4 says “exact”.

# If it breaks

FROM OFFICIAL DOCS, NOT TESTED BY US We did not hit these ourselves in our three runs. They are the common cases, with what to do.

## It gives no links at all

The chat may not be able to search the web. Look for a web search switch in the chat’s settings or under the message box. Without search, treat every number, quote and name as unverified and look it up yourself.

## The link opens a “page not found”

Count that as a made-up source. Reply:

```
That link does not open. Give me a working link to the original source, or
say you cannot find one.
```

### The page exists but the number is not on it

Reply:

```
I can't find that number on the page. Quote the exact sentence it comes from.
```

Then search the page for that sentence. If it is not there, drop the number.

### It says “you’re right, sorry” and gives a new number

Be more careful, not less. Check the new one the same way.

## Sources

Checked on 13 November 2026.

- Anthropic help: why Claude can give incorrect or misleading responses
- Anthropic help: enable and use web search
- The study from run 1: Waage and others (2021), BMC Nursing · free full text on PMC
- The press release from runs 2 and 3 (Norwegian), Fritt Ord via NTB

## Changelog

- 13 Nov 2026: First version, after three test runs in the Claude chat app.

## Report a broken step

Something did not work? Email us or DM @botsgowild. We check it, fix it and log the fix in the changelog.

---