



Never hand it over

Three tests, run in about ten seconds. Pass all three and you can move fast on everything else without second-guessing — which is really the point of knowing exactly where the line sits.



README

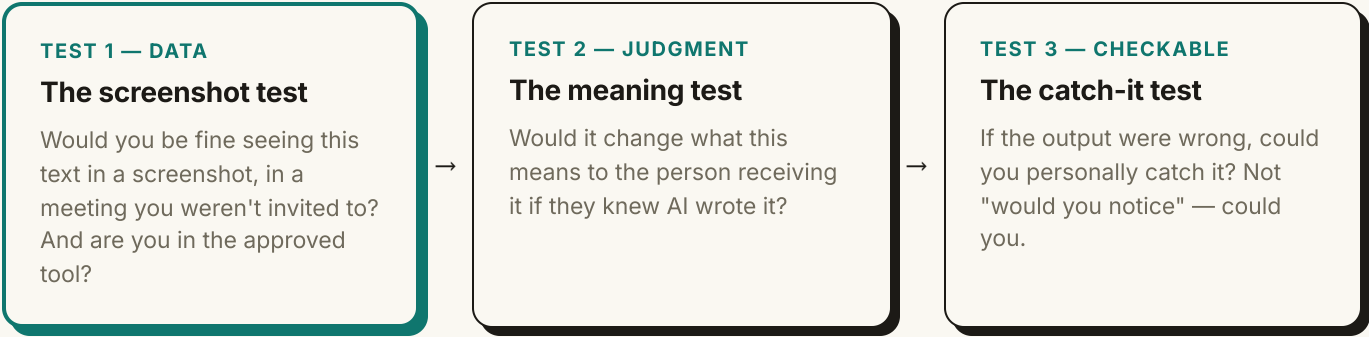
What this is. The data test, the judgment test, and the checkable test — plus prompts for stripping sensitive detail and forcing output you can actually verify.

What this isn't. "AI is dangerous, be careful." That isn't useful and you've heard it. These are three specific categories with a clear test each. Everything outside them is fair game.

The running example. Throughout, you're an operations manager at a mid-size company. Marcus is on your team. Support tickets need synthesising and a vendor comparison needs checking.

Three questions before you paste

Pass all three and hand it over. Fail one and do that part yourself — not the whole task, just the part that failed.



- a** **Test one is about which tool, not about AI.** The version your company approved handles data differently from the free one you signed up for with a personal email. Same logo, completely different agreement underneath. If you don't know which you're in, that's this week's task.
- b** **Test three gets more important as models improve, not less.** The output keeps getting more plausible. Your ability to verify it doesn't automatically keep up.

Where each test bites

TASK	FAILS	DO THIS INSTEAD
Customer records, salary bands	Test 1	Strip identifiers, keep the structure.
Performance review, hard feedback	Test 2	Write it; use AI only to tighten your words.
Analysis in a domain you don't know	Test 3	Force receipts, or get someone who can check.
Meeting notes into actions	Passes	Hand it over without a second thought.

Two ways to pass a failed test

Failing a test rarely means abandoning the task. Usually it means changing what you paste, or changing what you demand back.

1 · When test one fails — strip it first

COPY-PASTE · REDACT, KEEP THE SHAPE

I want help with the text below, but I need to remove identifying detail first. Rewrite it replacing:

- every person's name with Person A, Person B, and so on
- company and product names with Company A, Product A
- account numbers, emails and phone numbers with [redacted]
- exact figures with realistic round numbers, keeping the scale

Keep the structure, the tone and the problem exactly as they are. Then list anything you weren't sure about so I can check it myself.

Text: [paste]

2 · When test three fails — demand receipts

COPY-PASTE · MAKE IT CHECKABLE

Do the analysis below, but make every claim checkable.

For each conclusion, name the specific rows, columns or lines you used to reach it, so I can verify it without redoing the work.

Where the source doesn't actually support a conclusion, say "not supported by this data" instead of filling the gap. Separate what the data shows from what you're inferring, and label the inferences.

Data: [paste]

The ten-second `preflight`

Run this on anything you're unsure about. It's deliberately blunt, because a hedged answer here is worth nothing.

3 · Check before you commit

`COPY-PASTE · THE PREFLIGHT`

I'm about to hand you this task: [describe it in one line, no confidential detail]

Before I paste anything, assess it against three tests:

1. Data – is there likely to be material here I shouldn't paste into a general assistant?
2. Judgment – would the recipient read this differently if they knew AI wrote it?
3. Checkable – if you got this wrong, would someone in my position be able to catch it?

For each, answer pass or fail and say why in one line. If a test fails, tell me the smallest change that makes it pass.

The ten-minute task that beats everyone

`DO THIS ONCE`

Read your company's actual AI policy, properly, not skimmed. Most people are either ignoring it or quietly avoiding AI entirely. Knowing exactly where the line sits means you can work right up to it with confidence.

Three tasks, three verdicts

One week, one person. Each task fails a different test, and none of them get abandoned — they get adjusted.

Synthesise 400 support tickets

FAILS TEST 1 — DATA

The export has customer names and email addresses in it. The task itself is fine; the paste isn't. Run the redaction prompt first, then hand over the stripped version. Themes don't need real names to be accurate.

Write Marcus's performance review

FAILS TEST 2 — JUDGMENT

He'd read it completely differently if he knew. Write it yourself. Using AI to tighten a paragraph you already wrote and already mean is fine — asking it what you think of someone is not. The line is whether the judgment originated with you.

Compare two vendor contracts you don't fully understand

FAILS TEST 3 — CHECKABLE

The output will be articulate and confident, and you couldn't tell if a key clause were misread. Run the receipts prompt so every claim points at a specific clause — then you're verifying, not trusting.

Notice none of the three ended in "don't use AI for this." Two changed the input, one changed the demand. Only the review stayed entirely yours.

Three ways people get this wrong

01 Deciding it's all too risky

Not safe. Just slow.

These three categories are narrow on purpose. Everything outside them is fair game, and the whole reason to know the line precisely is so you can stop second-guessing the ninety percent of your work that was never near it.

02 Assuming approval covers judgment

Two different questions, routinely confused.

Your company approving a tool answers the data question. It says nothing about whether you should be delegating a decision that's genuinely yours, and nothing about whether you could catch an error if there were one.

03 Treating confidence as evidence

The failure that actually gets people burned.

Someone asks for analysis in an area they don't know, gets back something articulate, and forwards it because it reads well. It can sound completely certain and be completely wrong — a skill some coworkers have also mastered.

RECAP

You now have

- 1 Three tests — screenshot, meaning, catch-it — and a table of where each one bites.
- 2 A redaction prompt for test one and a receipts prompt for test three.
- 3 A blunt preflight you can run on anything you're unsure about.
- 4 Three worked verdicts where the task got adjusted rather than abandoned.

briandoai.com/resources