

A 60 MINUTE SESSION

Check what AI writes

A session any team lead can run, with no trainer in the room. Your team brings real work and practises checking what AI hands back: the answer, the sources it cites, and the changes it makes to your files.



YOUR TEAM LEAVES WITH THREE CHECKING HABITS

01

Spot what could be invented

Names, numbers, dates, quotes, sources and links.

02

Check it outside the chat

At the source itself, not on the AI's word.

03

Ask what was checked

In a cited report, and in a file an AI changed.

NEW IN THE OCTOBER 2026 EDITION

Exercise 3 checks what AI did, not just what it wrote: pairs trace a cited research report to its sources, and find what an AI changed in a shared file through its version history. Plus a practice sheet for teams that work with numbers, and a habit check for two weeks later.

Facts checked on 16 October 2026.

60 MINUTES

BEST WITH 4 TO 16 PEOPLE

ANY AI TOOL

NO TRAINER NEEDED

What this session does

AI tools write fluent, confident answers. They write a likely answer, and likely isn't the same as true. Now they also hand back cited reports, and some change shared files themselves. In one hour your team practises checking all three, on their own work.

WHO IT IS FOR

- Teams of 4 to 16 people who use AI for real work: emails, reports, summaries, research, spreadsheets.
- The team lead who runs it. You don't need to be an AI expert: this guide says what to do, what to say and what to watch for, minute by minute.
- Any tool. The habits are the same in any AI tool your organisation allows.

WHAT IT IS NOT

Not a course on one tool, and not a policy. If your organisation has rules about which AI tools to use and what may go into them, those come first. This session works inside them.

WHY NOW

- 50%

In Microsoft's 2026 Work Trend Index (20,000 people who use AI at work, in 10 countries), quality control of AI output topped the list of human skills that matter more as AI takes on work: 50% chose it.
- 66%

In the University of Melbourne and KPMG global study (more than 48,000 people, 47 countries, 2024 to 2025), in its section on AI at work, 66% said they rely on AI output without evaluating its accuracy.

And work now arrives as cited research reports, as edits made straight into shared files, and as whole tasks an AI carries out in your apps and browser, asking before important actions (as of Oct 2026). Sources: page 21.

WHAT'S IN THE BOX

FILE	WHAT IT IS	WHEN YOU USE IT
01 Facilitator guide A4 and US Letter	The plan, what to say, the exercises, every practice sheet with its answers, and the emails	Read it once before. Keep it open during
02 Slides PowerPoint and PDF	15 slides, one idea each, with speaker notes	On the screen
03 Participant handout A4 and US Letter	One page: the three checking habits, the routine, and space to write	Print one per person
04 Practice sheets A4 and US Letter	Four made-up sheets: a memo, a cited report, an edited file, and a table of figures. No answers on them	Print sheets 1 to 3 for each pair. Sheet 4 for teams that work with numbers
05 Emails a text file	The invite, the follow-up, a habit check for two weeks later, and the feedback question	A week before, the day after, two weeks after
06 Feedback form Word, A4 and US Letter	One question, on one page	At the close

THE FREE LICENCE

Free from Neosaipien under the free licence in LICENCE.txt. Run it with as many teams as you like, change the files to fit, and hand them to your own people, keeping the credit and the date.

Before the session

ONE WEEK BEFORE

- ☐ Book 60 minutes and a room with a screen. Online works too: see below.
- ☐ Send the invite. The one thing it asks: bring one real task.
- ☐ Check which AI tools your organisation allows, and what may be pasted into them. Say so in the invite. Page 4 shows where Claude, ChatGPT and Gemini keep what you'll check.

THE DAY BEFORE

- ☐ Print the handout, one per person, and practice sheets 1, 2 and 3, one set per pair. If your team works with numbers, add sheet 4.
- ☐ Try Exercise 1 yourself, on a task of your own. It takes ten minutes, and it gives you an example of your own to share.
- ☐ Optional, for part B of Exercise 3: set up a shared file an AI has edited, so the room can open its real version history (page 13).
- ☐ Open the slides on the screen you'll use. The speaker notes match this guide.

IN THE ROOM

- ☐ A laptop or phone with the AI tool, for each person or each pair
- ☐ Pens, a highlighter if you have one, and a timer
- ☐ A whiteboard or flip chart for the team habit

RUNNING IT ONLINE

Share the slides on screen. For the exercises, put people in pairs in breakout rooms, and send the handout and the practice sheets as PDFs the day before. Keep a timer where everyone can see it.

THE INVITE, TO COPY AND CHANGE

Subject: 60 minutes on checking what AI writes

Hi all,

On [day] at [time] we're spending an hour on one practical skill: checking what AI hands us before it goes anywhere. That's what it writes, the sources it cites, and the changes it makes to our files.

Please bring one real task you'd normally use AI for: an email to write, a document to summarise, or a question to research. If you can, also bring a document you're allowed to paste into [our AI tool], such as a report, a policy or meeting notes.

Bring a laptop or phone with [our AI tool] open. Nothing else to prepare, and nobody will be tested.

[Your name]

WHAT COUNTS AS A REAL TASK

- **Good:** something due this week that someone else will read.
- **Too small:** a quick fact, like the date of a holiday.
- **Too sensitive:** anything you're not allowed to paste into the tool. Use the practice sheets instead.

YOUR JOB IN THE ROOM

Keep the clock. The exercises are where people learn, so protect their time.

You don't need to know more than the room. If someone knows a lot about AI, ask them to help a pair.

Ask, don't answer. When someone asks whether something is right, ask where they'd check it.

Have a backup. If the AI tool is down, pairs share a laptop or use the practice sheets.

Your team's AI tool, side by side

This box covers Claude, ChatGPT and Gemini. The three habits are the same in all three; only the buttons differ. Use the tool your organisation approves for work, and find its column before the session, so you can show the room where to look. Features reach accounts in waves: if a button is missing, it may be the plan, the country or an admin setting.

AS OF OCT 2026	CLAUDE	CHATGPT	GEMINI
Ask, or hand over a task	One conversation for both. Since 16 Sep 2026, Cowork is part of ordinary Claude chat	Chat for questions; Work for tasks and finished files, in one app since 9 Jul 2026	Chat; Task for agent work, on personal accounts with Google AI Pro or Ultra. At work, Gemini sits inside Gmail, Docs and Sheets
A cited report (Exercise 3, part A)	Turn on web search, then + and Research	Type /Deepresearch, or + and Deep research . Review its plan first	In the text box, Add Files , then Deep Research . Edit the plan, then Start research
A file it edited (part B)	Claude for Excel cites the cells it used and warns before it overwrites. Claude Docs and Slides in any chat	Add-ins for Excel, Word, PowerPoint and Google Sheets. Work makes documents and decks	Gemini in Docs, Sheets and Slides, for Google Workspace users
Before it acts	Asks before taking an action, by default. Connectors: Needs approval	Asks you to approve important actions. Apps: Always ask	Confirms some actions and sign-ins. Write actions in connected apps always need your yes
For software only	Claude Code	Codex	Antigravity

WORDS PEOPLE WILL ASK ABOUT

- **Model:** the brain you pick from a menu. A stronger model still guesses, so check it the same way.
- **Project:** a folder of files and instructions the AI works from.
- **Connector:** a signed-in link from the AI to another app, such as your files or calendar. ChatGPT calls it an app; Gemini calls it a connected app.
- **MCP:** the shared standard those links are built on, like a USB-C port for AI apps. The AI reaches only what your own account can reach.
- **Skill:** a saved how-to. All three now read the same SKILL.md format.

WHICH MODEL FOR A CAREFUL CHECK

- **Claude:** Sonnet 5.5 every day; Opus 5.5 for a careful report.
- **ChatGPT:** GPT-6 at Medium every day; High for a careful report.
- **Gemini:** Flash every day; Pro, thinking High, for a careful report.

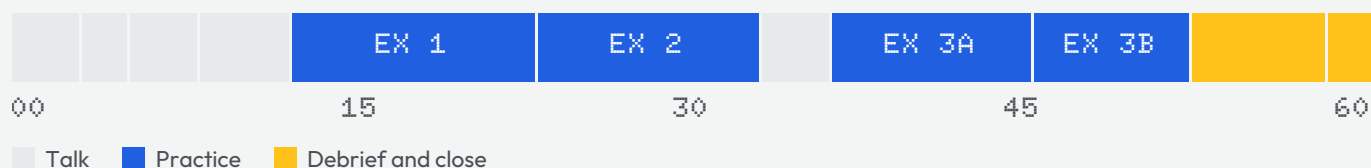
Start on the default. A bigger model writes better guesses, not checked facts, so the habits stay the same.

MY RECOMMENDATION

Claude first, ChatGPT second, Gemini third. Claude for the simplest start to finish: one conversation, asking before it acts, finished files on every plan. ChatGPT for its Work agent, and for research where you control the sources. Gemini where your company runs on Google's apps, and for notes in Google Meet. Your organisation's approved tool comes first, whatever I recommend.

The 60 minutes

Talk takes about 15 minutes, practice 37, and the wrap-up 8. The practice is where the learning is. Every timing here is an estimate: adjust it to your room.



STARTS	ABOUT, MIN	WHAT HAPPENS	SLIDES	PAGE
00:00	3	Welcome, and what this hour is for	1	6
00:03	2	Why it matters now: the 2026 picture	2	6
00:05	3	The classic case, 2023: six cases that didn't exist	3 to 5	7
00:08	4	Why it happens: a likely answer, not a looked-up one	6 and 7	7
00:12	11	Exercise 1: Spot what could be invented	8	8
00:23	10	Exercise 2: Check a summary against its source	9	8
00:33	3	Check what it did, not just what it wrote	10 and 11	11
00:36	9	Exercise 3, part A: Trace a cited report	12	12
00:45	7	Exercise 3, part B: Find what an AI changed	13	12
00:52	6	Debrief	14	17
00:58	2	One team habit, and close	15	17

KEEPING TIME

- If you fall behind, cut talk, not practice.
- Read each exercise's steps aloud, then start the timer.
- Give a one minute warning before each exercise ends.

ONLY 45 MINUTES?

Give one figure on slide 2 and skip the maths story on slide 10. Take 2 minutes off Exercises 1 and 2. Run one part of Exercise 3: A for reports, B if AI edits your shared files. Keep the debrief to two questions. The slides show the 60 minute times, so say the shorter ones aloud.

HAVE 90 MINUTES?

Add sheet 4, the numbers sheet, as an extra exercise (about 10 minutes). Give Exercises 1, 2 and 3 four more minutes each, so pairs can share with the room, and the debrief 5 more.

HOW TO READ THE NOTES

SAY

Points to make in your own words. This isn't a script to read. Where words are in bold, use those words: those are the facts, and they have been checked.

DO

What you and the group do.

WATCH FOR

What might go wrong, and how to steer it.

KEEP IT ACCURATE

The limits of the facts. Stay inside them, even when someone asks for more.

00:00 ABOUT 3 MIN / SLIDE 1

Welcome

SAY

- This hour is about one habit: checking what AI hands you before it goes anywhere. That's what it writes, the sources it cites, and the changes it makes to your files.
- We'll practise it on your own work, in three habits: spot what could be invented, check it outside the chat, ask what was checked.
- Nobody is being tested. Anything wrong we find is the point.

DO

- Ask everyone to open their real task and their AI tool now, so it's ready for Exercise 1.
- Hand out the handout. Point to the boxes at the bottom: they fill them in as they go.

00:03 ABOUT 2 MIN / SLIDE 2

Why it matters now

SAY

- Checking came top. **In Microsoft's 2026 Work Trend Index, a survey of 20,000 people who use AI at work in 10 countries, quality control of AI output topped the list of human skills that matter more as AI takes on work: 50% chose it.**
- **In a global study by the University of Melbourne and KPMG, in its section on AI at work, 66% said they rely on AI output without evaluating its accuracy, and 56% have made mistakes in their work because of AI.**
- And the work looks different now. AI tools hand back cited research reports, and some edit your shared files themselves. We'll practise checking both.
- Then lead into slide 3: first, the classic case, from 2023.

KEEP IT ACCURATE

- The Work Trend Index asked which human skills matter more as AI takes on more work. Two topped the list: quality control of AI output (50%), then critical thinking (46%). The survey was run for Microsoft by Edelman Data x Intelligence, February to April 2026, and published on 5 May 2026.
- Microsoft sells AI tools, and the figures are what people reported. Say **in this survey**, not **everyone**.
- The Melbourne and KPMG study surveyed more than 48,000 people in 47 countries, between November 2024 and January 2025. The 66% and 56% are from its section on AI at work.

WHY THE 2026 PICTURE COMES FIRST

The tools have changed a lot since 2023: as of Oct 2026 they write cited reports, edit shared files, and carry out whole tasks in your apps and browser. The habit hasn't changed. Starting with this year's figures tells the room the session is about the work they do now, and the classic case then shows what the habit is for.

GOT A STORY OF YOUR OWN?

If your team has its own example of AI getting something wrong at work, tell it after the classic case (slide 5), without naming anyone. A story from your own work means more to the room than one from outside it.

00:05 ABOUT 3 MIN / SLIDES 3 TO 5

The classic case, 2023

SAY

■ SLIDE 3

In 2023, a lawyer asked ChatGPT for court cases to support his argument. It gave him six that looked real, with names, dates and quotes from the judges.

■ SLIDE 4

They went to a federal judge, and not one of them existed.

■ SLIDE 5

When he asked whether the cases were real, he was asking the same machine. It said yes.

DO

- After slide 5, ask the room: **What would you have checked, and where?** Take two answers. Say you'll come back to the question at the end.

IF THEY ASK

- **Was it real?** Yes. It's in the court's own order, which is public.
- **Isn't that old news?** Yes, it's from 2023, and the tools have changed a lot since. The habit it teaches hasn't: whatever the tool, the check happens outside the chat. That's why it's the classic case.

KEEP IT ACCURATE

- Say **a lawyer** and **a federal judge**. Don't name anyone, the law firm or the case.
- Six is the number of invented cases, not everything ChatGPT gave him. Never say **only six** or **all six**.
- Don't say when he asked whether the cases were real, or why: the court record doesn't settle either. Don't say he never checked, or add what happened to him afterwards.
- Source: court order on sanctions, U.S. District Court, Southern District of New York, 22 June 2023. A story about checking work: nothing here is legal advice.

00:08 ABOUT 4 MIN / SLIDES 6 AND 7

Why it happens

SAY

■ SLIDE 6

The AI wasn't lying. It was guessing. It writes a likely answer, one word at a time. It knows what a court case looks like: two names with a v in the middle, a court, a year, and a confident quote. So it wrote a likely version of that. **And likely isn't the same as true.**

■ SLIDE 7

Asking it, are you sure, isn't checking. It's the same machine vouching for its own answer. Check outside the chat: the original document, a source you trust, or a person free to disagree.

- Why those six kinds? That's where likely and true part company most often, and where a mistake does the most harm. Exercise 1 starts there.

WATCH FOR

- **"So AI is useless"** Steer it: it's good at likely. The job is to check the parts where likely isn't good enough: names, numbers, dates, quotes, sources and links.

KEEP IT ACCURATE

- "Wasn't lying, was guessing" is our way of putting it, based on how the companies that make these models describe them. It isn't what the court said.
- Don't say "AI can't check anything" or "never ask again" to the room. A fresh chat, or a second look against a source you paste in, is a different thing from "are you sure?" in the same chat.

00:12 ABOUT 11 MIN / SLIDE 8 / EXERCISE 1

Spot what could be invented

Goal: everyone finds the parts of a real answer that could be invented, and checks two of them outside the chat.

STEPS, ON THE SLIDE

- 1 **2 MIN** **Run your task** in your AI tool, or use an answer it gave you before.
- 2 **3 MIN** **Underline** every name, number, date, quote, source and link.
- 3 **2 MIN** **Mark each one.** K: I know it's right. S: I know the source to check it. ?: I can't tell.
- 4 **3 MIN** **Check two**, the two that would hurt most if they were wrong. Check them outside the chat.
- 5 **1 MIN** **Tell your neighbour** one thing you found.

SAY

- Before step 4: **Don't ask the AI to check itself. Go to the source.**
- After step 5, to the room: **Who found something they couldn't check? What did you do?**

WATCH FOR

- **Only numbers get underlined.** Prompt: names of people, products and policies, quotes, dates, links. Quotes and links can be invented too.
- **Nothing turns out wrong.** A good result: they know which parts were checked, and that's the habit.
- **Lots of question marks.** Ask: who would know? Often that person is the check.

NO AI TOOL, OR A CONFIDENTIAL TASK?

Use practice sheet 1 (page 9). Do steps 2 and 3 on the summary only, without reading the memo yet; Exercise 2 checks it. Its marked example also shows a room that's unsure how to mark.

00:23 ABOUT 10 MIN / SLIDE 9 / EXERCISE 2

Check a summary against its source

Goal: everyone holds one summary against its source and finds what's missing, changed or added.

STEPS, ON THE SLIDE

- 1 **2 MIN** **Get a summary** of about 100 words of a document you're allowed to paste into your AI tool, or use practice sheet 1.
- 2 **5 MIN** **Read it beside the source.** Mark each difference. M: missing. C: changed. A: added.
- 3 **3 MIN** **Rewrite** the one sentence that would cause the most trouble if it went out as it is. Then compare with your neighbour.

SAY

- **Don't ask: is this right? Ask: what's missing, what's changed, what's added?**
- **Read for what it leaves out, not only for what it says.**

WATCH FOR

- **Skimming.** Ask pairs to go line by line: one reads the summary aloud, the other follows in the source.
- **Added is the hardest to see.** A sentence that sounds sensible but isn't in the source counts as added. So does a quote you can't find in the source.

THE PRACTICE SHEETS

Sheet 1 has six planted mistakes, two of each kind; the answers are on page 10, for you only. A team that works with numbers uses sheet 4, the ticket figures (page 19), instead: the same steps, and every number recalculated from the table.

The desk move

PRACTICE TEXT, MADE UP FOR THIS WORKSHOP

Exercise 1: on the summary only, underline every name, number, date, quote, source and link, and mark each K, S or ?.
Exercise 2: now read the memo. Mark each difference M (missing), C (changed) or A (added).

HOW TO MARK: AN EXAMPLE, FROM A DIFFERENT MADE-UP ANSWER

From 1 May **K**, all expenses go through the new expenses tool. Receipts must be uploaded within 30 days **S**, and claims over 500 **S** need a manager's approval. The finance team says "paper receipts are a thing of the past" **?** and its guide is on the intranet under Finance, Expenses **S**.

THE SOURCE: A MEMO

To: everyone on the third floor
From: Facilities
Subject: Moving to the fifth floor

We're moving to the fifth floor on 9 March. The new floor has 40 desks for the 55 of us, so most desks will be booked through the desk app, a day at a time.

You can keep a fixed desk if you're in the office four or more days a week and your manager agrees. Ask your manager by 20 February.

Lockers on the third floor must be emptied by 6 March. Anything left after that will be cleared.

The kitchen and both meeting rooms on the fifth floor open on the first day. The quiet room opens in April.

If you need a desk set up for a health reason, tell Facilities before 27 February and it will be ready on the first day.

THE SUMMARY: SIX PLANTED MISTAKES

The team moves to the fifth floor on 2 March. The new floor has 55 desks, and most are booked through the desk app, a day at a time. Anyone in the office four or more days a week can keep a fixed desk. The kitchen and both meeting rooms open on the first day, and the quiet room opens in April. Facilities says the move "will make it easier for everyone to find a desk" and the new layout is expected to cut office costs by a third. If you need a desk set up for a health reason, tell Facilities before 27 February, and it will be ready on the first day.

M MISSINGC CHANGEDA ADDED

THE SENTENCE WE'D REWRITE FIRST

HOW WE'D REWRITE IT

FOR YOU, NOT THE ROOM

Practice sheet 1: the answers

Six planted mistakes, two of each kind. Every other sentence in the summary matches the memo.

#	KIND	IN THE SUMMARY	IN THE MEMO
1	C	on 2 March	on 9 March
2	C	The new floor has 55 desks	40 desks for the 55 of us
3	M	Anyone in the office four or more days a week can keep a fixed desk	Only if their manager agrees, and they must ask by 20 February. A pair may count these as two: that's right too.
4	M	Nothing about lockers	Lockers must be emptied by 6 March, or they'll be cleared.
5	A	The quote "will make it easier for everyone to find a desk"	No such words. A quote you can't find in the source counts as added.
6	A	expected to cut office costs by a third	Not in the memo at all.

These match the memo: most desks are booked through the app a day at a time; the kitchen and both meeting rooms open on the first day; the quiet room opens in April; a desk for a health reason means telling Facilities before 27 February, and it will be ready on the first day.

Which to rewrite first? Any answer is fair if the pair can say what it would cost. Point out mistakes 3 and 4: nothing in the summary reads wrong there, which is why missing things are the hardest to see.

IF A PAIR FINDS SOMETHING ELSE

Ask them to show you the two sentences side by side. If the meaning really differs, they're right and this key is wrong: tell the room, and use it as the lesson. A checker's job is to find what the key missed, too.

00:33 ABOUT 3 MIN / SLIDES 10 AND 11

Check what it did, not just what it wrote

SAY

■ SLIDE 10

One more story, from this year. In September 2026, an AI turned OpenAI's proof on a famous problem about how water flows into code that a computer can check, and the computer check passed. **So three mathematicians laid that proof's code beside its paper, to see if the two matched. They didn't. In at least two places, the code went its own way. In one, it proved less than the paper claimed. Which means the computer had checked a different proof of the same result, not the one on paper. They're not saying the written proof is wrong, only that the pass can't tell you if it's right.**

■ SLIDE 11

So when an AI tells you its work passed a check, ask what was checked. That's the third habit. AI work now comes in two shapes that need it.

- **A report with sources.** Deep research tools write a report with citations. The sources look like proof. Ask: does each source say what the report claims?
- **An edit to a shared file.** Some AI tools now change the file itself. Ask: what did it change, and was that asked for?

KEEP IT ACCURATE

- Their paper is a preprint, not peer reviewed. Keep that with it, and keep "they're not saying the written proof is wrong" before the last line.
- Don't say the proof or its result is wrong, unproven or solved. What passed was the code, and "proved less" is about one step, not the whole result.
- Say **three mathematicians**. Don't name them or call them experts.
- If someone mentions OpenAI's collection of 719 maths papers from October: that's a separate release. The mismatch was found in the September proof.
- Sources: OpenAI, 8 Sep 2026. Bastounis, Circelli and Hansen, preprint, arXiv 2610.08144, 6 Oct 2026.

WHAT THE TOOLS SAY, AS OF OCT 2026

- OpenAI says ChatGPT's deep research gives "a structured report with citations or source links so you can verify the information" (OpenAI Help Center).
- "Where supported and authorized, ChatGPT can update the source file directly" in Google Drive (OpenAI, 13 Aug 2026).
- Claude for Excel answers with cell references you can click, and warns before it overwrites a cell (Anthropic). Claude's Research and Gemini's Deep Research write cited reports too (page 4).

THE ROUTINE, FOR ANYTHING THAT MATTERS

SORT

Decide how much to check: who sees it, and can it be undone?

LIST

Write down every claim that can be checked.

TRACE

Find each one at its source, outside the chat.

HEDGE

Keep the words that make it true.

LOG

Record what was checked, the date, and when to check again.

The same routine, in the same words, as the AI Fact-Check Kit. Exercise 3 uses it on a report, then on a file.

00:36 ABOUT 9 MIN / SLIDE 12 / EXERCISE 3, PART A

Trace a cited report

Goal: every pair traces a cited report's claims to their sources, and finds the two that don't say what is claimed.

STEPS, ON THE SLIDE AND SHEET 2

- 1 **1 min.** **1 min.** Who will see this report, and what will be decided on it? Decide how much to check.
- 2 **1 min.** **1 min.** Number every claim that can be checked.
- 3 **4 min.** **4 min.** Open each source (the right-hand box) and mark each claim Y, P or N.
- 4 **1 min.** **1 min.** Where a claim is true, underline the words that keep it true.
- 5 **1 min.** **1 min.** Fill in one row of the log.
- 6 **1 min.** **1 min.** The report says its sources were checked. What was checked?

SAY

- **A source in square brackets isn't a check. It's a place to look.**
- **Trace means you read the source itself, not the AI's summary of it.**
- **Before step 6: When an AI tells you its work passed a check, ask what was checked.**

WATCH FOR

- **Stopping at the link.** Pairs see that a source exists and move on. Ask: but does it say that?
- **Fixing by guessing.** If a source doesn't say it, the fix is to say what the source does say, or to drop the claim.
- **The recommendation.** It isn't a fact to trace, but it rests on claims 2 and 3.

YOUR TEAM USES A DEEP RESEARCH TOOL?

Run the same six steps on one real report instead, with sources you're allowed to open. Trace three claims, not all of them. Sheet 2's answers are on page 16.

00:45 ABOUT 7 MIN / SLIDE 13 / EXERCISE 3, PART B

Find what an AI changed

Goal: every pair finds what an AI changed in a shared file, beyond what it was asked to do.

STEPS, ON THE SLIDE AND SHEET 3

- 1 **1 min.** **1 min.** Read the request. What was the AI asked to do?
- 2 **3 min.** **3 min.** Compare the two versions line by line. Mark each difference W, C, M or A.
- 3 **2 min.** **2 min.** Ask of each C, M or A: was it asked for? Would anyone have seen it without the version history?
- 4 **1 min.** **1 min.** Decide: keep, fix, or restore version 4. Who on our team checks AI edits?

SAY

- **This time it didn't write you an answer. It changed the file everyone works from.**
- **The version history is where you see what it did.**

WATCH FOR

- **Only the wording gets marked.** Point at line 1, and at the line that's gone.
- **Blaming the tool.** The request was loose and nobody checked. Both are fixable.

Part B with your own files

Sheet 3 works anywhere. If your team's AI tool can edit shared files, the real thing teaches more. Set it up the day before, and only in tools and files your organisation allows.

THE DAY BEFORE

- 1 Make a copy of a short, harmless document of your own, such as a how-to page, in a shared file.
- 2 Ask your approved AI tool to edit the copy, with this request: **Tidy the wording. Keep the meaning.**
- 3 Don't read the result. Share the copy with the room.

IN THE ROOM

- 1 Open the file's version history together.
- 2 Compare the version before the AI's edit with the one after it, and run steps 2 to 4 from sheet 3.
- 3 If nothing changed the meaning, say so. That's a result: now you know what was checked.

KEEP IT SAFE

- Use a copy, never the live file.
- Nothing confidential, and nothing with customers' or colleagues' details.
- Don't connect email, messaging or payment tools for this.

WHERE THE VERSION HISTORY IS, AS OF OCT 2026

- **Google Docs, Sheets and Slides:** at the top right, click **Last edit**, then pick an earlier version in the panel to see it or restore it. You need edit access to see earlier versions, and versions can be merged unless you name one first.
- **Word, Excel and PowerPoint in Microsoft 365:** click the title of your file, then **Version history**. On the web: File, then Version History. Version history "only works for files stored in OneDrive or SharePoint in Microsoft 365" (Microsoft Support).
- **Claude for Excel:** its answers point to the cells they used, it keeps formulas intact when it changes values, and it warns before it overwrites.

From Google's, Microsoft's and Anthropic's own help pages, read on 16 Oct 2026. Menus move: if yours differs, search your tool's help for **version history**.

WHY THIS MATTERS

When an AI edits a shared file, there may be no draft to read: the change goes straight into the file everyone uses. The version history is where you see what it did, so someone has to look. In Google's editors, name the current version before the AI edits, so it stays separate.

WHEN THE AI TOOK AN ACTION, AS OF OCT 2026

All three assistants now carry out whole tasks: Claude from any conversation, ChatGPT in Work, Gemini with Task (personal accounts). A sent email has no version history, so check before the action and after it.

- **Before:** keep approvals on for anything that sends, books, posts or deletes. Claude asks before taking an action by default; ChatGPT Work asks you to approve important actions; Gemini confirms some actions and sign-ins. For connected apps, choose **Needs approval** in Claude or **Always ask** in ChatGPT, never **Allow all actions**.
- **After:** read the steps it took, then look where it acted: the sent folder, the calendar, the channel. Whoever asked answers for it.

The cited report

PRACTICE TEXT, MADE UP FOR THIS WORKSHOP

An AI research tool wrote this report for a made-up company. The box on the right is what you find when you open each source. Run the routine, and mark each claim **Y** (the source says it), **P** (it says only part of it) or **N** (it doesn't say it).

SORT
Decide how much to check.

LIST
Write down every claim that can be checked.

TRACE
Find each one at its source, outside the chat.

HEDGE
Keep the words that make it true.

LOG
Record what was checked, the date, and when to check again.

THE REPORT

Research report, 15 October 2026
Question: should we offer next-day delivery?

FINDINGS

1

In our September survey, 61% of the customers who replied said delivery time mattered most when they chose a supplier [1].

☐

2

Most suppliers in our sector already offer next-day delivery [2].

☐

3

Our courier charges the same rate for next-day as for standard delivery on parcels up to 20 kg [3].

☐

4

Complaints about late delivery doubled from 2024 to 2025 [4].

☐

RECOMMENDATION

Offer next-day delivery on all parcels up to 20 kg from January.

SOURCES

[1] Customer survey, September 2026. Internal file.
[2] Sector newsletter, 2 October 2026. Web page.
[3] Courier rate card, valid from 1 July 2026. PDF.
[4] Complaints log, updated 5 January 2026. Spreadsheet.

Check: all four sources were found, and all four are dated 2026.

EACH SOURCE, OPENED

[1] CUSTOMER SURVEY

Question 4: what mattered most when you chose a supplier? Delivery time: 61% of the 410 customers who replied.

[2] SECTOR NEWSLETTER

Of the 12 suppliers we spoke to, 4 now offer next-day delivery, and 3 more plan to by next summer.

[3] COURIER RATE CARD

Next-day delivery: same rate as standard for parcels up to 5 kg. Over 5 kg, a next-day surcharge applies.

[4] COMPLAINTS LOG

Complaints about late delivery: 38 in 2024, 76 in 2025.

LOG ONE CLAIM YOU TRACED

CLAIM

SOURCE

DATE READ

VERDICT

HEDGE KEPT

CHECKED BY

Verdict: true, true with a hedge, wrong, or could not confirm.

ASK WHAT WAS CHECKED

The report's last line says all four sources were found and are dated 2026. What did that check, and what did it not?

PRACTICE TEXT, MADE UP FOR THIS WORKSHOP

VERSION 4 / BEFORE

- VERSION 5 / AFTER THE
AI'S EDIT

- ## VERSION HISTORY

14 Oct 2026, 09:20
Edited by the AI
assistant, at a
colleague's request

2 Sep 2026, 16:05
Edited by the returns
team lead

Tidy the wording. Keep the meaning.

WHICH CHANGES WERE
ASKED FOR?

[illegible]

KEEP, FIX, OR
RESTORE VERSION 4?
WHY?

[illegible]

WHO ON OUR TEAM CHECKS AI EDITS?

[illegible]

FOR YOU, NOT THE ROOM

Practice sheet 2: the answers

#	MARK	THE REPORT SAYS	THE SOURCE SAYS
1	Y	61% of the customers who replied said delivery time mattered most [1]	The same. The hedge to keep is of the customers who replied (410 of them), not of our customers .
2	N	Most suppliers in our sector already offer next-day delivery [2]	4 of the 12 suppliers the newsletter spoke to offer it, and 3 more plan to. That's a third, not most. P is fair if the pair says why.
3	N	The same rate for next-day as for standard on parcels up to 20 kg [3]	The same rate only up to 5 kg. Over 5 kg, a surcharge applies. P is fair too: true up to 5 kg, wrong as written.
4	Y	Complaints about late delivery doubled from 2024 to 2025 [4]	38 in 2024, 76 in 2025: exactly double.
R	–	Offer next-day delivery on all parcels up to 20 kg	Not a fact to trace, but it rests on claims 2 and 3. With both corrected, it needs rethinking.

What was checked? "All four sources were found, and all four are dated 2026" checks that the sources exist and are recent. It doesn't check that they say what the report claims. Two of them don't.

Practice sheet 3: the answers

#	MARK	WHAT CHANGED	ASKED FOR?
1	C	Line 1: 30 days became 14 days	No. The request was to keep the meaning.
2	W	In line 2 "The item ... its" became "Items ... their"	Yes: wording only.
3	W	In line 3 "go back to" became "are made to"	Yes: wording only.
4	M	Version 4's line 4, the rule "Refunds over 200 need a manager's sign-off" is gone	No. A rule disappeared, and nothing in version 5 reads wrong.
5	W	In version 5's line 4 (line 5 before) "can't ... it's" became "cannot ... it is"	Yes: wording only.
6	A	In the last line "who reply within one working day" was added	No. It's a promise nobody made.

Three of the six changes alter the meaning, and none of the three was asked for. **Decide:** restore version 4, or fix lines 1, 4 and 6 by hand, and tell the person who asked for the edit. Then agree who checks AI edits to shared files, and when.

THE POINT TO LAND

Neither the report nor the file looked wrong. The report had sources and a check line; the file read better than before. Both needed someone to ask what was checked.

00:52 ABOUT 6 MIN / SLIDE 14

Debrief

Pick three of the four questions. Go round the room, or take answers from anyone.

- 1 What did you find today that you'd have missed last week?
- 2 Which parts of our work need checking every time, and which are fine with a quick read?
- 3 Where do we check things outside the chat: which document, which system, which person?
- 4 When AI hands us a report or changes our files, who checks it, and how?

THEN

- Go back to the classic case and ask again: **What would you have checked, and where?** Ask whether the answers have changed since the start.

WATCH FOR

- **A rule for everything.** Checking every word isn't the aim. Checking the parts that matter is: that's what Sort is for.

00:58 ABOUT 2 MIN / SLIDE 15

One team habit, and close

DO

- Agree one sentence. Type it into the white box on slide 15, or write it on the board. Everyone copies it onto their handout.
- If the room is stuck, offer one of these:

Anything with a name, number, date or quote gets checked against its source before it leaves the team.

Every cited report gets its sources traced before we decide on it.

When AI edits a shared file, whoever asked for it checks the version history the same day.

SAY

- The last line on the slide: **Check outside the chat.**
- You'll send the habit and the handout by email the next working day, and check in on it in two weeks.
- Hand out the feedback form, or ask its one question aloud.

QUESTIONS PEOPLE ASK

Can I paste our documents into AI?

Only into tools your organisation has approved for that kind of information. If you're not sure, use the practice sheets today and ask whoever looks after IT or data.

Which AI tool is best for this?

This session doesn't pick one. The habits work the same in any of them.

Our AI doesn't edit files. Is part B for us?

Yes. Any shared file someone pasted AI text into has the same risk, and the same fix: compare it with the version before.

What if nothing it wrote was wrong?

Then you know, and you know which parts you checked. That's the habit.

Won't checking take longer than doing it myself?

You don't check everything. Sort first, then check the parts that would hurt if they were wrong.

Who checks when the AI did the work?

Whoever asked for it, because "The AI did it" is never the answer to "who checked this?"

After the session

- ☐ The next working day, send the follow-up below with the handout attached.
- ☐ Put the team habit somewhere people will see it.
- ☐ If your organisation keeps a record of AI training, log this session in it. The Team AI Policy Kit has one.
- ☐ Two weeks later, send the habit check, and take five minutes at a team meeting on what people found.
- ☐ All four emails are also in the box as a text file, ready to paste: the invite, the follow-up, the habit check and the feedback question.

THE FOLLOW-UP, THE NEXT WORKING DAY

Subject: Checking what AI writes: our habit, and the one-pager

Hi all,

Thanks for [day]'s session. Here's what we agreed, and the one-page handout is attached.

Our team habit: [the sentence you agreed]

The three checking habits

1. Spot what could be invented: names, numbers, dates, quotes, sources and links.
2. Check it outside the chat: at the source itself. What's missing, what's changed, what's added?
3. Ask what was checked: in a cited report, does each source say what's claimed? In a file an AI edited, what does the version history show?

For anything that matters: Sort, List, Trace, Hedge, Log.

One thing to stop: asking "are you sure?" It's the same machine vouching for its own answer.

One ask: reply with one thing you checked this week and what you found. I'll share a few, without names, at our next team meeting.

[Your name]

THE HABIT CHECK, TWO WEEKS LATER

Subject: Two weeks on: is our AI checking habit holding?

Hi all,

Two weeks ago we agreed: [the sentence you agreed]

A quick check, one line each, by [day]:

1. One thing you checked since the session, and what you found.
2. Did the habit hold? If not, what got in the way?

At [team meeting] we'll spend five minutes on what you found, and keep the habit, change it, or add a second one.

[Your name]

THE FEEDBACK QUESTION

The form in the box asks one thing. Hand it out at the close, or ask it by email:

"What is one thing you will check differently after this session?"

Read the answers before the habit check: they show whether the habit landed.

RUNNING IT ACROSS MANY TEAMS

- **Train the leads first.** Run the session once with the team leads as the room. Each lead then runs it with their own team, from this guide.
- **Keep it the same.** Every team gets the same slides, practice sheets and four debrief questions, so you can compare their answers.
- **Report what changed.** Collect three things from each team: the habit it agreed, the feedback form answers, and the replies to the habit check. Report how many teams say their habit held after two weeks, and a few things people found, without names.

The free licence covers as many teams as you like. Change the slides and sheets to fit, and keep the credit and the date on them.

The ticket figures

PRACTICE FIGURES, MADE UP FOR THIS WORKSHOP

An AI assistant wrote a summary of a made-up support team's quarter from the table below. The summary has five planted mistakes, and some sentences that are right. Recalculate every number from the table, then mark each sentence **OK**, **C** (changed or miscalculated) or **A** (added), and write **M** beside anything the summary leaves out.

THE SOURCE: THE TEAM'S TABLE

SUPPORT TICKETS	JUL	AUG	SEP*
North	310	290	335
South	420	445	470
East	180	205	190
West	250	240	265
Total	1,160	1,180	1,260
Average days to resolve a ticket	2.6	2.4	2.2

* September figures are provisional until 10 October.

YOUR WORKING

[illegible]

THE SENTENCE THAT WOULD DO THE
MOST HARM

THE SUMMARY

Q3 in brief, written by an AI assistant from the table

- 1 The team handled 3,800 tickets in the quarter. ☐
- 2 September was the busiest month, up 12% on August. ☐
- 3 South had the most tickets in every month, about 37% of the quarter. ☐
- 4 North's tickets rose every month. ☐
- 5 East was the only region where tickets fell in September. ☐
- 6 The average time to resolve a ticket fell every month, to 2.2 days in September. ☐
- 7 At this rate, the team will pass 4,000 tickets in the fourth quarter. ☐

OK RIGHT C CHANGED A ADDED

M MISSING

HOW WE'D REWRITE IT

FOR YOU, NOT THE ROOM

Practice sheet 4: the answers

Use sheet 4 in place of sheet 1 in Exercise 2, for a team that works with numbers, or on a second run of the session. Same steps, with one more: recalculate every number from the table. Five planted mistakes; sentences 3, 5 and 6 are right.

#	MARK	THE SUMMARY SAYS	THE TABLE SAYS
1	C	3,800 tickets in the quarter	3,600: 1,160 + 1,180 + 1,260.
2	C	September was the busiest month, up 12% on August	Busiest, yes. But up about 7%: 1,260 against 1,180 is 80 more.
3	OK	South had the most tickets every month, about 37% of the quarter	Right: 1,335 of 3,600 is 37%.
4	C	North's tickets rose every month	No: 310 in July, 290 in August, then 335.
5	OK	East was the only region where tickets fell in September	Right: East went from 205 to 190; the rest rose.
6	OK	Average time to resolve fell every month, to 2.2 days	Right: 2.6, 2.4, 2.2.
7	A	The team will pass 4,000 tickets in the fourth quarter	No forecast in the table. Added.
-	M	Nothing about the footnote	September's figures are provisional until 10 October, so the summary's September numbers aren't final.
-	M	Nothing about West	Fair too: West is in the table but nowhere in the summary.

THE POINT TO LAND

The right sentences make the wrong ones believable. A summary that's mostly right is the one people stop checking.

WHICH TO REWRITE FIRST?

Sentence 1 or 7 is a good pick: a total or a forecast is what gets copied into a report. Any answer is fair if the pair can say what it would cost.

Checked against

Every fact in this box was read at its source on 16 October 2026, or checked against its primary source on the date shown. The two OpenAI lines are the exception: OpenAI's help pages refused a direct read that day, so both were read in the Internet Archive's copies, dated below. Company pages change without notice, so every line about a product is marked as of Oct 2026. Nothing in this box is legal advice. The practice sheets and every example are made up.

The court case. Opinion and Order on Sanctions, U.S. District Court, Southern District of New York, 22 June 2023. Read 7 Oct 2026.

How models write, and why they guess. OpenAI, *Why language models hallucinate*, 5 Sep 2025. Anthropic, *Tracing the thoughts of a large language model*, 27 Mar 2025. Read 7 Oct 2026.

"Are you sure?" isn't checking. Studies of assistants challenged on their answers: Sharma and others, 2023; Laban and others, 2023; Fanous and others, 2025; Kim and Khashabi, 2025. Anthropic, *Reduce hallucinations*. Read 7 Oct 2026.

Quotes and links can be invented too. Anthropic, *Reduce hallucinations*. Columbia Journalism Review, Tow Center, a test of AI search tools, 6 Mar 2025. Read 4 Oct 2026.

The 2026 Work Trend Index. Microsoft, 5 May 2026; 20,000 AI users in 10 countries, surveyed by Edelman Data x Intelligence, Feb to Apr 2026. Read 16 Oct 2026.

microsoft.com/en-us/worklab/work-trend-index/agents-human-agency-and-the-opportunity-for-every-organization

66% and 56%. University of Melbourne with KPMG, *Trust, attitudes and use of Artificial Intelligence: A global study 2025*, more than 48,000 people, 47 countries, Nov 2024 to Jan 2025. KPMG's page read 16 Oct 2026; the university's media release of 30 Apr 2025 in an archived copy.

kpmg.com/xx/en/our-insights/ai-and-technology/trust-attitudes-and-use-of-ai.html

Deep research. OpenAI Help Center, *Deep research in ChatGPT*. OpenAI's help pages refused a direct read on 16 Oct 2026, so this was read in the Internet Archive's copy of 3 Oct 2026.

help.openai.com/en/articles/10500283-deep-research-in-chatgpt

ChatGPT updating Drive files. OpenAI, ChatGPT release notes, entry of 13 Aug 2026, read in the Internet Archive's copy of 10 Oct 2026.

help.openai.com/en/articles/6825453-chatgpt-release-notes

Claude for Excel. Anthropic, *Use Claude for Excel*. Read 16 Oct 2026.

claude.com/docs/office-agents/excel.md

The three tools side by side, and when the AI took an action. The vendors' own pages, as of Oct 2026, read 16 Oct 2026, among them:

Anthropic, *Claude Cowork and chat are now one Claude*, 16 Sep 2026; OpenAI, *ChatGPT Work and Codex*; Google, *Manage your tasks and workflows in Gemini Apps*; the Model Context Protocol's own introduction.

claude.com/resources/articles/cowork-is-now-claude

help.openai.com/en/articles/20001275-chatgpt-work-and-codex

support.google.com/gemini/answer/17094507

modelcontextprotocol.io/docs/getting-started/intro

Version history. Google Docs Editors Help, *Find what's changed in a file*, and Microsoft Support, *View previous versions of Office files*. Both read 16 Oct 2026.

support.google.com/docs/answer/190843

support.microsoft.com/en-us/office/view-previous-versions-of-office-files-5c1e076f-a9c9-41b8-8ace-f77b9642e2c2

The maths proof. OpenAI's post on its Navier-Stokes proof, 8 Sep 2026. Bastounis, Circelli and Hansen, *Navier-Stokes lost in translation*, a preprint, not peer reviewed, version 1 of 6 Oct 2026. Read 14 Oct 2026, and still version 1 on 16 Oct 2026.

arxiv.org/abs/2610.08144

WHEN TO CHECK AGAIN

The product lines on pages 2, 4, 6, 11 and 13, by 16 Nov 2026. The "are you sure" line, by 4 Jan 2027. The maths preprint, whenever it changes version or is published. Checked on the vendors' pages on 16 Oct 2026: these products change monthly, so check the settings in your own tool.

After this session

Brief it, check it, keep it safe: this session is the Check it part of the three habits. Here's where to go next.

THE SECOND SESSION

Workshop in a Box: Brief an Agent

The same format, 60 minutes: hand work to an AI agent with a six-part brief (Goal / Done looks like / Context / May touch, must not / Limits / Stop and ask), and decide how far to let it go.

FOR EACH MANAGER

The AI Fact-Check Kit

The same routine, Sort, List, Trace, Hedge, Log, for one person: review cards for each type of AI work, a claims log and a practice pack.

FOR YOUR RULES

The Team AI Policy Kit

A policy, a staff guide and a record of AI training. Log this session there.

FOR THE WHOLE ORGANISATION

The Team Pack

Every kit and both workshops in one free download, run as a 90-day programme.

FREE: THE FIELD GUIDE

Level 2, Everyday skills: checking answers, research with sources, and more, at neosaipien.com/learn/

LIVE, WITH YOUR TEAM

In the Team workshop I run this live with your team, in half a day, on its own real work:
neosaipien.com/services/#workshops

HOW THIS BOX WAS CHECKED

Every fact in it was checked against its primary source, and each source is listed with the date it was read on page 21. If a company changes its pages, the "as of" date tells you how old a line is.

Book a call

A 20 minute call about your team and its AI work. Neosaipien is AI training and consulting for the people doing the work: the Leadership briefing, Executive coaching, the Team workshop, Workflow automation and AI implementation.

neosaipien.com / samson@neosaipien.com