

The AI Bootcamp

20

How to Tell Real Research From Recall

In this chapter you'll learn

How to tell a researched answer from a recalled one, what it cannot see, and a 4-minute check on any sourced report

The AI Bootcamp

THE PROBLEM

Most AI education is a pile of disconnected tips. A LinkedIn carousel here, a YouTube tutorial there, a \$2,000 cohort course that goes out of date before the refund window closes. None of it is built to be read in order, so nothing compounds. You finish twelve videos and can recite twelve facts, not perform one new skill.

THE FIX

The AI Bootcamp is built the other way. 90 skills, packed 2 to 3 per chapter, 36 chapters, 3 tracks. Every chapter states what it needs from you and what it hands to the chapter after it, so skill 47 only works because skill 12 already landed. Read it in order and by the end you are not collecting facts about AI, you are running work through it.

LETTER FROM THE EDITOR

We built this because we kept meeting smart people who felt behind on AI for no reason other than nobody had organized the material for them.

This book does not assume you are technical, and it does not water anything down either. Every chapter earns its place: a real skill, a real exercise, a one-page companion you can flip back to in month three when you have forgotten the details but still need the answer.

If a chapter loses you, or a chapter changes how you work, I want to hear about it either way.

Kris

Editor, AI Central

Write to me at editor@thecentral.ai

THE 3 TRACKS

AI 101

Get running

Pick a tool, set it up, write a clear instruction, use it safely at work.

AI 202

Prompt properly

The framework behind every good instruction, and catching AI when it is wrong.

AI 303

Build without watching it

Connect AI to your tools, and automations that run while you are not looking.

IN THIS LESSON YOU'LL LEARN

By the end of this chapter you can tell whether an answer came off the live web or out of training, name the 4 categories of material your assistant structurally cannot reach, use a deep research mode on the jobs it is actually good at, and check a sourced research output in about 4 minutes.

Research is the first thing most senior people hand to AI, and the place a wrong answer travels furthest, because a sourced report gets forwarded. The failure mode is not a tool lying loudly. It is an articulate, well-structured, fully cited document resting on sources nobody opened.

CONTENTS	20.1	Two very different answers come out of the same box
	20.2	The web it can reach is smaller than the web
	20.3	The long mode is a different tool, not a better one
	20.4	Treat it as work from a bright, unaccountable junior

TOPICS REQUIRED

Chapter 19, and the habit it left you with of checking the number yourself rather than trusting the table it arrived in. Bring one real research question from your work: the kind you would currently hand to an analyst or lose an afternoon to. Chapter 8's verification habit is the floor this chapter builds on.

AFTER THIS CHAPTER

Chapter 21 assumes you are holding a piece of research you half believe, and that you are willing to have the assistant attack it rather than agree with it.

20.1 Two very different answers come out of the same box

One is written from what the model absorbed in training, months or years ago. The other is written from pages it fetched thirty seconds ago. They arrive in the same font, in the same tone, and only one of them can be checked.

Recalled	Researched
Written from what the model absorbed during training. No page was opened to produce it. TRAITS · Arrives fast, usually in seconds · No links, or links written from memory · Frozen at the training cutoff PROS + Fine for stable knowledge, definitions, method + Fast enough to use in the middle of a thought CONS – Cannot know anything after its cutoff – A web address it produces from memory may not exist	Written from pages fetched while the answer was being written. You can see which ones. TRAITS · Slower, and usually shows what it searched for · Carries clickable citations next to claims · Can reach something published this morning PROS + Every claim has an address you can open + Current, within whatever the index has picked up CONS – Only reaches pages that are public and crawlable – A citation existing is not proof it supports the claim

Fig 20.1 Both answers are written in the same confident voice and land in the same place. The only reliable tell is whether there is something you can click.

20.1

The tell. Clickable citations. No links means it answered from memory, whatever it sounded like.

Ask plainly. "Search the web for this and cite your sources" works in all four assistants.

Mixed answers are normal.

One reply can be half fetched and half recalled. The uncited half is the recalled half.

Here is the problem this chapter exists to fix: your assistant produces two completely different kinds of answer, in the same box, in the same voice, and it very rarely tells you which one you just received.

The first kind is recall. The model writes from what it absorbed during training, which is a compressed impression of a very large amount of text that stopped at a fixed date. Nothing was looked up. For a definition, a method, a piece of settled history, that is perfectly good and it is fast. For anything that moves, it is a confident description of the past.

The second kind is retrieval. The assistant runs actual searches, opens actual pages, and writes from what it just read. All four assistants can do this now, and each has its own logic for deciding when to bother. Claude has had web search on every plan globally since May 2025. Gemini leans on Google's index by default. Copilot reaches the web and, at work, your files and mail. ChatGPT searches when the question looks like it needs it, or when you say so.

Which means the practical question is not which assistant searches. It is how you tell, on any given answer, whether it did. The tell is citations you can click. Most assistants also show a line while they work, naming what they searched for, and it is worth watching that line rather than scrolling past it. What you must not do is read confidence as evidence of research: the two are unrelated.

This matters most in the cases where you would never think to check. Ask about a regulation, a competitor's pricing, a person's current role, a product's current limits, and a recalled answer will describe a version of the world that was true when the model was trained. It will not tell you that. It has no way of knowing what it does not know.

Open a chat and ask the same question twice: once plainly, once with the words "search the web and cite your sources". Read both answers side by side. The gap between them is the whole of this section.

20.2

The web it can reach is smaller than the web

The most expensive assumption a senior professional makes about AI research is that a tool connected to the internet has the internet. It has the free, public, crawlable, already-indexed slice of it.

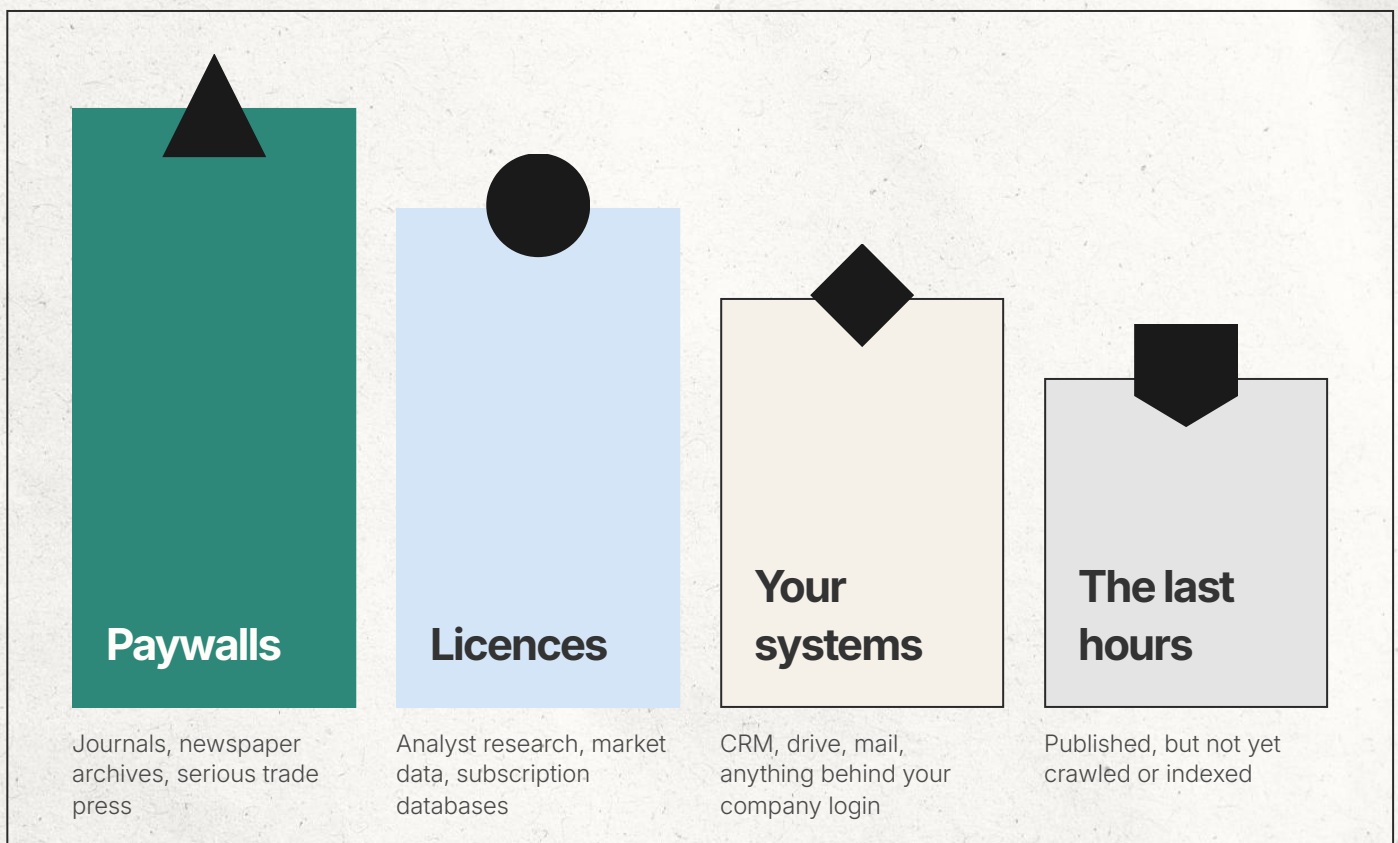


Fig 20.2 Four walls, and not one of them is announced in the answer. The bar heights are visual rhythm only, not a measure of how much material sits behind each wall.

20.2

Blocked on purpose. From 1 July 2025 Cloudflare began blocking AI crawlers by default across the sites it fronts. A great deal of publishing is now off limits by policy, not by accident.

Abstract, not paper. It reads the free preview of a paywalled piece and writes as though it read the whole thing.

The best follow-up. "What could you not reach, and what would change your answer?" is the highest-value question in this chapter.

Your assistant does not fail at the wall. It writes straight through it, in the same voice, and the answer contains no marker where the evidence stopped.

Start with paywalls, because they cover most of what a professional would actually cite. Academic journals, newspaper archives, the serious trade press in your sector: the assistant fetches the public page, which is a title, an abstract, two paragraphs and a subscribe button. It then produces a paragraph that reads like a summary of the research. Sometimes it is right, because the abstract said so. Sometimes it is an inference from a headline.

Then there is material that is blocked deliberately. Publishers have been closing their doors to AI crawlers for two years, and since July 2025 Cloudflare has blocked them by default on the sites it sits in front of. The awkward part is that the blocks are not always respected: the Tow Center's early-2025 tests found tools correctly answering questions about articles from publishers that had explicitly disallowed them. So you cannot infer from a source appearing that it was fairly reached, and you cannot infer from a source missing that it does not exist.

Licensed industry data is a harder wall still. Analyst subscriptions, market data terminals, licensed panel data, the paid database your team already pays for: none of it is on the open web, so none of it is in the answer. What you get instead is the press release that quotes one number from it, or a vendor blog citing a figure from three years ago. That is a real difference in quality and the report will not flag it.

Your own systems are invisible by default too. Your CRM, your drive, your mail, last quarter's board pack: unless you have deliberately connected them, they do not exist as far as the assistant is concerned. Connecting them is a genuine option and it changes the answer completely, which is why Chapter 26 covers connectors and Chapter 27 covers what it means to give something permission to reach into your work.

Last, the clock. Something published this morning may be hours away from being findable, and an assistant asked about it will either say nothing or fall back on recall. For anything breaking, treat AI research as a tool for context, not for the newest fact. Run one question from your own field and read the source list for who is not on it.

20.3 The long mode is a different tool, not a better one

All four assistants now have a mode that goes away for several minutes, reads dozens of pages and returns a long report full of citations. It is genuinely useful, and it is routinely spent on the wrong questions.

Search-backed reply	Deep research run
The assistant runs a handful of searches inside a normal answer and replies in seconds. TRAITS · Seconds, inside the conversation · A few sources, usually the obvious ones · Costs nothing beyond a normal message PROS + Fast enough to stay in the flow of the work + Easy to challenge and rerun immediately CONS – Stops as soon as it has enough to answer – Tends to miss the source that disagrees	A separate mode that plans, searches repeatedly, reads dozens of pages, then writes a structured report. TRAITS · Roughly 5 to 30 minutes, unattended · Dozens of sources and a bibliography · Metered: a small monthly allowance on most plans PROS + Maps ground you do not already know + Surfaces disagreement, not just consensus CONS – More citations also means more bad citations – Long and authoritative-looking is not the same as right

Fig 20.3 Choosing between these is a question about the shape of the answer you need. The long mode is not a better search, it is a survey, and surveys are wasted on questions you can already frame.

20.3

ChatGPT Deep research. Roughly 5 to 30 minutes. A small monthly allowance free, larger on paid plans.

Claude Research. Paid plans only. Searches the web, plus your Google Workspace files if you connect them.

Gemini Deep Research. Shows you an editable plan first, then usually 5 to 10 minutes of work.

Copilot Researcher, under Agents. Reaches your work files, mail and meetings as well as the web. Paid licence.

You have a question that would cost an analyst two days. That is what this mode was built for, and most people spend theirs on questions a single search would have closed in nine seconds.

Mechanically, all four do the same thing. The assistant turns your question into a plan, fans out across many searches, opens and reads the pages it thinks matter, then writes a structured document with a bibliography. Gemini shows you that plan before it starts, which is the most useful feature in the category and the one people click past fastest. Edit it. Twenty seconds spent narrowing the plan is worth more than any prompt wording.

The costs are real and they are the reason to be deliberate. Runs take roughly 5 to 30 minutes depending on the tool and the question. They are metered: most plans give you a modest monthly allowance rather than unlimited use, and Claude's version and Copilot's require a paid plan at all. You are choosing to spend one of a small number of these, so spend it on something worth 20 minutes of waiting.

What it is genuinely good at is unfamiliar ground. A market you have never sold into, a regulation that just landed in your inbox, a vendor category with eleven players you cannot name, a scientific question adjacent to your field. In those cases you do not yet know the right search terms, and the report's real product is not its conclusion. It is a map and a reading list: the names, the factions, the terminology, the four sources you should now read yourself.

Where it is oversold is your own field. The assistant has no sense of your profession's tacit hierarchy of evidence. It weighs a well-optimised content marketing page and a regulator's own notice at roughly the same value, because on the open web they look similar. Where you know the field, you already carry that hierarchy in your head, and you will usually beat the machine on judgement even while it beats you on volume.

There is one counterintuitive finding worth carrying. Because these modes read more, people assume their citations are safer. Work published in April 2026 across more than 221,000 cited links found the opposite: deep research agents produce more citations per question than ordinary search-backed chat, and hallucinate them at a higher rate. More footnotes is not more rigour. It is more surface area for the check in the next section.

20.4

Treat it as work from a bright, unaccountable junior

The output is fast, articulate, well-organised and carries no consequence for being wrong. You already know how to read work like that. The only new muscle is that the citations have to be opened rather than trusted.

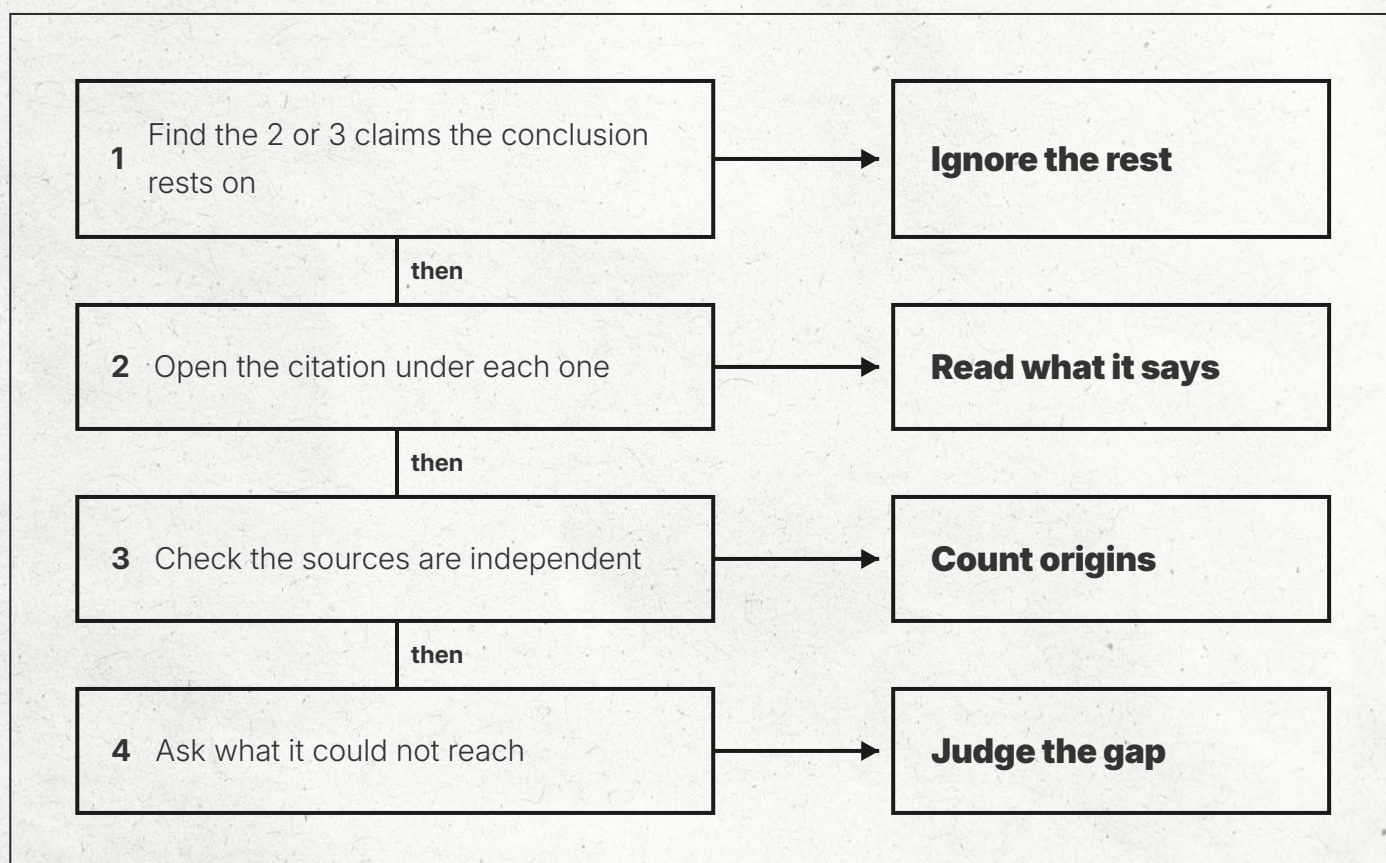


Fig 20.4 About 4 minutes, in this order. Step 2 is the one everybody skips and the one that catches the most.

20.4

Open, do not trust. Across 221,000 cited links studied in 2026, 3 to 13% were fabricated outright and 5 to 18% did not resolve at all.

Five sources, one origin. Syndicated copies of one wire story are one source. Tools cite the copies, not the original.

Chapter 8 is the floor. This is the research-specific version of the verification habit you built there.

You would not put a junior analyst's memo in front of your board without opening the two numbers the recommendation turns on. Read an AI research output exactly that way, with one addition: the junior cannot be asked why, and will not be embarrassed next quarter.

The citation is where this is won or lost, and a citation can fail in four separate ways. The link is dead. The link was invented and never existed. The link is real but about a different thing. Or, most often and most dangerously, the link is real and relevant and simply does not say what the sentence above it claims. The first two are measurable: research published in April 2026 across more than 221,000 cited links found 3 to 13% fabricated outright and 5 to 18% failing to resolve. The fourth kind is not measurable from your armchair. You have to open it.

The second failure is subtler and it is the one that catches experienced people. A confident synthesis with five citations can rest entirely on one weak source, because the other four either repeat it or address something adjacent. This is not hypothetical: the Tow Center's early-2025 test of 1,600 queries across eight AI search tools found them citing syndicated copies rather than originals as a matter of course, and wrong more than 60% of the time overall. The same study found the tools almost never admit doubt, one of them flagging low confidence 15 times in 200 answers and never once declining. So count origins, not links.

Then ask the question the report will never volunteer: what did you not find, what is behind a paywall or a subscription, and which of these claims rests on a single source. A good assistant will answer that honestly and the answer is frequently more useful than the report. It tells you where the real work still is.

Some research you should not hand over at all. Anything where the answer depends on knowing which sources are respected and which are marketing. Anything regulated, contractual or legally consequential, where a fabricated citation is not an inconvenience but a professional problem. Anything where the cost of being wrong lands on someone other than you. In those cases the assistant is still useful for the reading list and the vocabulary. The judgement stays with a person who knows the field.

The next research output that reaches you, from any source, run the four steps with a timer. It takes about 4 minutes, and you will find something in the first one.

READING ENDS HERE

Everything before was reading. Everything after is doing.

3 exercises, About 20 minutes and one real decision that opens Chapter 21.

Do these before Chapter 21

3 exercises. About 20 minutes

Ask your assistant a question in your own field you already know the answer to, telling it to search the web and cite its sources. Open every citation. Write how many actually said what the answer claimed they said.

Run one real research question from your work. Fill both boxes from what the answer shows you, not from what you were hoping for.

SOURCES IT ACTUALLY USED

WHAT IT COULD NOT REACH

Take the last research output you were handed by anyone, human or machine. Find the 2 claims the conclusion rests on, write them here, and go and check them.

Chapter 20 on one page

RECALLED VS RESEARCHED

No clickable citations means it answered from training, not from the web. Force the difference with "search the web and cite your sources"

THE 4 BLIND SPOTS

Paywalled journals and archives · licensed analyst and market data · anything behind your company login · anything published in the last few hours

THE 4-MINUTE CHECK

Find the 2 or 3 load-bearing claims · open the citation under each · count independent origins, not links · ask what it could not reach

DEEP RESEARCH MODES

ChatGPT deep research · Claude Research (paid) · Gemini Deep Research (edit the plan first) · Copilot Researcher (paid, also reaches work files). Minutes long, metered, best on ground you do not already know

THE BEST FOLLOW-UP

"What did you not find, what is behind a paywall, and which of these claims rests on one source?"

BRING TO CHAPTER 21

The research output you checked, and the list of what your assistant could not reach. Chapter 21 turns the same assistant against the conclusions you have just formed.

BOTTOM LINE

You can tell a researched answer from a recalled one, name what your assistant structurally cannot see, and check a sourced report in about 4 minutes.

NEXT CHAPTER

Get AI to Argue With You

Chapter 21 cures the agreeableness. It covers how to get a real objection instead of a compliment, how to red-team your own work through it, and what a useful disagreement actually looks like on screen.

BRING WITH YOU

- ☐ A research output you have already checked
- ☐ One conclusion of your own you would rather not be wrong about

The Ultimate AI Library

explore the library



WHO MAKES THIS



AI Central Media is a London-based editorial team. Our flagship publication covers practical AI for senior professionals, and 100+ AI companies, SaaS platforms and growth teams have advertised with us since 2023.

100+

AI companies, SaaS platforms, education brands and growth teams have advertised with us since 2023

613K

Accounts reached every month, across the 7 channels we operate

151

Countries our readers are in, plus all 50 US states

PROUD PARTNER OF

GAMMA

IIElevenLabs

guidde.

Outskill

Luma AI

HubSpot

attio

S Synthflow

Notion

Replit

We turn attention into pipeline.

visit thecentral.ai

