



Using AI Tools Effectively at Work

AI at work 01 | The tool families as of 2026 and what each is for, task first then tool, five levels of delegation, centaur and cyborg ways of working, the jagged frontier in daily use, a should-I-use-AI decision, session habits and a worked task

#AIatWork #Tools #Delegation #JaggedFrontier

Think of it as a new colleague

A new hire on day one	An AI assistant
Widely read, fast, eager	Trained on a huge text corpus, answers in seconds
Does not know your clients or history	Knows only what is in the chat and attached files
Needs a clear brief to do good work	Needs goal, audience, material and format
Rarely admits not knowing	Answers confidently even when guessing
You check the first deliverables closely	You verify in proportion to the stakes
Gets better with feedback in the conversation	Improves with follow-up instructions in the same chat
WHERE THE ANALOGY BREAKS A new hire learns and remembers your feedback for good; an assistant forgets it in a new chat unless you save it as project instructions, and it does not get more reliable with time on its own.	

An AI assistant behaves like a brilliant new hire on day one

Prerequisite in one line

An LLM predicts plausible text; fluent output is not checked output (00-welcome 02, responsible-ai 06). Everything on this sheet follows from that.

The tool families as of 2026

Tool	Draft and rewrite	Your documents	Web with links	Inside your apps	Multi-step tasks
ASK AND DRAFT					
General assistant ChatGPT, Claude, Gemini, Copilot, Le Chat	✓	✓	✓	⚠	⚠
Embedded assistant Microsoft 365 Copilot, Gemini in Workspace	✓	✓	⚠	✓	⚠
FIND AND RESEARCH					
AI search Perplexity, assistants' search modes	⚠	⚠	✓	✗	⚠
Deep research mode in several assistants	✓	⚠	✓	✗	✓
Source-grounded notebook NotebookLM	⚠	✓	✗	✗	✗
SPECIALISED					
Meeting assistant Teams, Zoom, Meet built-ins; notetakers	⚠	⚠	✗	✓	✗
Writing and translation DeepL, Grammarly	✓	⚠	✗	✓	✗
Coding assistant or agent GitHub Copilot, Cursor, Claude Code, Codex	✓	✓	⚠	✓	✓
✓ yes / full ⚠ partial ✗ no / none					

What each tool family typically does as of 2026 (✓ core, ~ some products or modes, ✗ not its job). Products change monthly: check the features your organisation has licensed

Task	Best fit	Trap
Reply to a long mail thread	embedded assistant in the mail client	it may miss a commitment buried in an attachment
Summarise a 60-page report	general assistant or notebook with the file attached	a summary of the file, not of the web: attach it
Answer a factual question	AI search with links	an assistant without search answers from memory
Market or literature scan	deep research mode	long reports hide weak sources; open the key ones
Questions over your own document set	source-grounded notebook	answers stop at what you uploaded

Spreadsheet formula, SQL, regex	general or embedded assistant	test on rows whose answer you know
Minutes and action items	meeting assistant	names and owners often wrong; consent to record

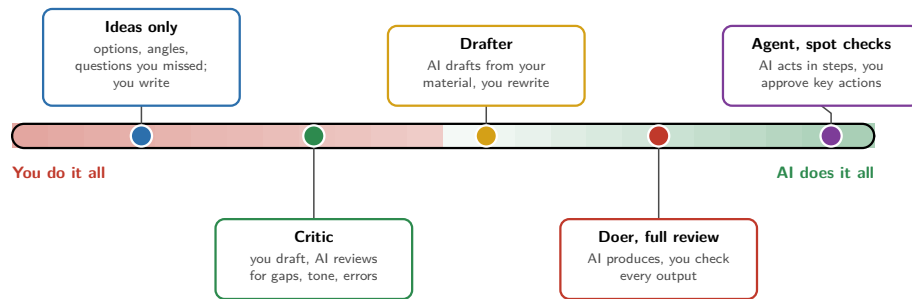
Task first, then tool: the usual best fit and the trap to avoid

💡 Set up a project per recurring job

For any task you repeat (weekly report, client X, grant applications) create a project or workspace with standing instructions (audience, tone, format, what to avoid) and reference files. It saves re-briefing every chat and keeps outputs consistent.

● How much to delegate

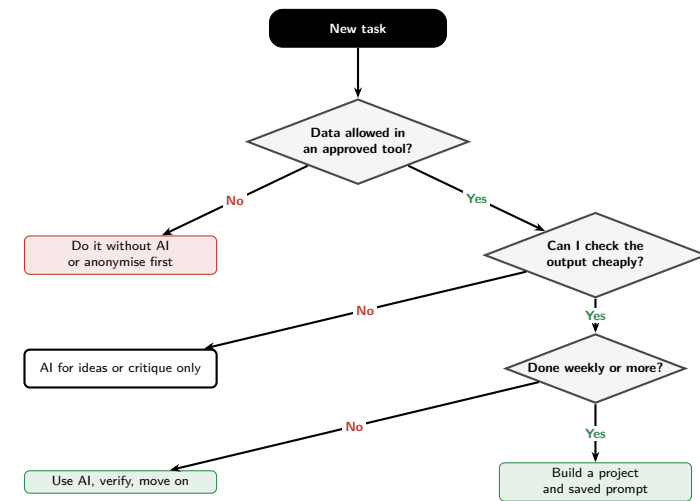
“How much of the task the AI does”



Five delegation levels. Move right only when the output is cheap to check or the stakes are low; move left when you cannot verify it

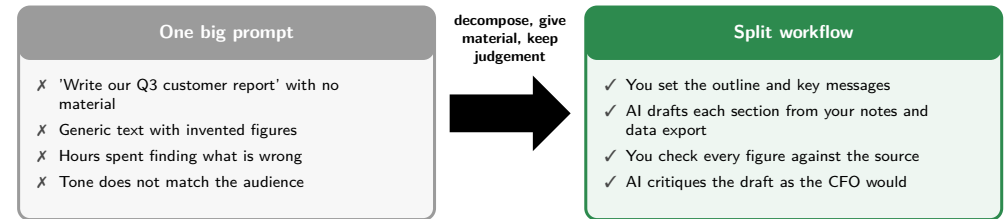


Two ways of working observed in the 2023 consultant study, and when each fits



Should I use AI for this task, and how?

🔄 From pasting the task to designing the workflow



The same quarterly report, done by pasting the whole task vs by splitting it

PROBLEM Summarise 300 customer comments into themes for a leadership meeting. Baseline: about 4 h by hand. Is AI worth it, and how?

GIVEN

- Comments contain customer names and order numbers
- Approved enterprise assistant available; consumer tools not allowed for customer data
- Stakes: medium (decisions on next quarter's priorities)

1 **Classify the task**

summarise material you provide

Why: inside the frontier for current assistants

2 **Check the data rule**

customer data → approved tool only

Why: sheet 05 traffic light: amber

3 **Pick the level**

Drafter: AI proposes themes with counts and example quotes

Why: themes are cheap to check by sampling

4 **Plan the check**

read 30 random comments, confirm each lands in a theme; verify every quote verbatim

Why: catches missed themes and invented quotes

5 **Estimate time**

20 min prompting + 40 min sample check + 30 min editing = 1.5 h

Why: include checking time, not only generation

ANSWER Worth it: about 2.5 h saved on this run (4 h – 1.5 h), if the sample check passes; log the actual times

CHECK

- ✓ If the sample check finds a missed theme, add a step: ask for comments that fit no theme, then re-read those

Choosing tool, delegation level and check for one real task (times are this example's assumptions, to be replaced by your own log, sheet 06)

✓ Session habits

Every AI session

✓ baseline / mandatory □ recommended – verify in your context

- ✓ Use an approved tool for the data you will paste
- ✓ One chat per task; start fresh when the topic changes long mixed chats drift and carry stale context
- **MUST** State the deliverable: who reads it, length, format
- **MUST** Attach the source material instead of describing it
- **SHOULD** Ask for uncertainty: 'mark anything you are not sure of' helps, but does not replace your check
- **SHOULD** Decide the check before you read the output
- **COULD** Save prompts that worked to your library

MYTH

- ✗ Always use the newest, biggest model
- ✗ Try every new tool
- ✗ If AI can do it, delegate it fully
- ✗ Chat history is my memory

FACT

- ✓ Use the one your data rules allow, then the cheapest that does the job
Reasoning modes are slower; for a short rewrite they add wait, not quality.
- ✓ Depth with two or three tools beats breadth
Switching costs time and spreads data across more vendors.
- ✓ Delegate only what you can check
Full delegation of an unverifiable output moves the risk to you.
- ✓ Save what matters in your own files
Chats get long, are hard to search and may be deleted by retention rules.

Habits that look efficient and are not

Key takeaways

01

Task first, tool second

general, embedded, search, notebook, coding

02

Pick a delegation level

ideas, critic, drafter, doer, agent

03

Delegate what you can check

the cost of checking sets the level

04

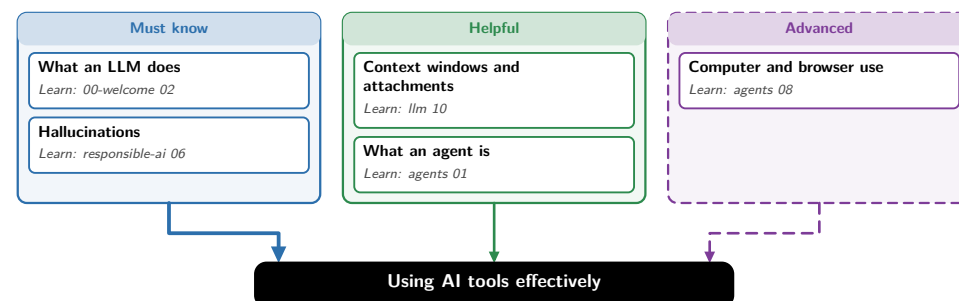
Split big tasks

outline and judgement stay with you

05

Projects for repeated work

standing instructions and files



What this sheet builds on, and where to go deeper

Q1 Name five AI tool families used at work as of 2026.

A General assistants, embedded assistants, AI search and deep research, source-grounded notebooks, coding assistants (also meeting and writing tools).

Q2 Best tool kind for a factual question?

A AI search with linked sources; an assistant without search answers from memory.

Q3 What is a source-grounded notebook good for?

A Answering questions only from documents you upload, with a citation to the passage (e.g. NotebookLM).

Q4 List the five delegation levels.

A Ideas only, critic, drafter, doer with full review, agent with spot checks.

Q5 What sets how far you can delegate?

A How cheaply you can check the output, and what an undetected error would cost.

Q6 Centaur vs cyborg?

A Centaur splits the task between person and AI; cyborg interleaves them at every step (Dell'Acqua et al. 2023).

Q7 Why one chat per task?

A Long mixed chats drift and carry stale or irrelevant context into new answers.

Q8 Why set up a project or workspace?

A Standing instructions and reference files are reused, so you stop re-briefing and outputs stay consistent.

Q9 In the worked example, what made the time estimate honest?

A It counted prompting, sample checking and editing time, not only generation time.