

30 / 60 / 90 DAYS

The AI Rollout Plan

From licences bought to work that changed. A 90-day plan for the people who run the organisation.

DAYS 1 TO 30

Decide and set the rules

DAYS 31 TO 60

Pilot

DAYS 61 TO 90

Scale, fix or stop

INSIDE

- The plan, stage by stage, with the worksheet or product for each step
- Which platform first: Claude, ChatGPT or Gemini, and the order to roll them out in
- A plan on a page, a 90-day checklist and a leadership kickoff deck
- A job picker and measures tracker, and a monthly update template
- Questions for IT or your admin, and the rules leaders should know

Inside, and how to use it

The tools are bought. A few people are keen. There are no shared rules, the leaders are not aligned, and nothing measures whether anything got better. This plan fixes those three things in 90 days.

THE PLAN

03	Why now
04	The leader's three jobs
05	The 90 days at a glance
06	Choosing the platform: Claude, ChatGPT, Gemini
09	Days 1 to 30: decide and set the rules
16	Days 31 to 60: pilot
18	Days 61 to 90: scale, fix or stop
21	After day 90: the monthly rhythm
22	Measuring: time saved is not value
23	Cost
24	Rules for leaders
25	Questions for IT or your admin
26	Checked against
27	What next

THE FILES

- **The plan** you are reading, A4 and US Letter.
- **Plan on a page**, a PDF to pin up and an editable PowerPoint file.
- **Job picker and measures tracker**, an Excel workbook, with the same sheets as CSV files.
- **Leadership kickoff deck**, 9 slides with speaker notes, PowerPoint and PDF.
- **The 90-day checklist**, one page, PDF and Word.
- **Monthly leadership update**, a one-page Word template.
- **Leadership alignment worksheet**, one page, PDF.
- **Questions for IT or your admin**, PDF and Word.
- **Outfit**, the free font the Office files use, with its licence.
- **LICENCE.txt**, the free licence: what you may do with all of these files.

HOW TO USE IT

01

Read pages 3 to 8 before anything else: about 25 minutes.

03

Book the leadership session for week 1, with the kickoff deck and the worksheet.

02

Name an owner and give them this pack (page 9).

04

Work down the checklist. Each step says who, when and what to use.

WHO DOES WHAT

The leader	Decides, sets the rules, uses AI on real work, presents at day 90.	About 1 hour a week
The leadership team	The session, the policy, a decision for each job, the monthly update.	3 meetings, then monthly
The owner	Runs the plan: register, IT questions, reviews, training record, updates.	2 to 4 hours a week
Two pilot managers	Score the jobs, take the baseline, run the weekly review.	About 1 hour a week
IT or the admin	Answers the questions, sets “Ask first”, keeps the logs, can pull the plug.	A few hours, weeks 1 to 4
Everyone	Reads the guide, says what they use, reports what goes wrong.	20 minutes, once

Examples: every example is made up and names no real organisation or person. **Times:** every time is an estimate. **Rules and law:** plain-English orientation, not legal advice. The steps name other Neosaipien products where one fits, and say what to do without them. Neosaipien is not affiliated with Anthropic, OpenAI or Google.

The tools are not the hard part

This year's large surveys point the same way: the organisation around people moves results more than the tools or the individuals do.

2x

Organisational factors such as culture, manager support and talent practices “account for more than 2x the reported AI impact of individual factors” (67% vs 32%).

Microsoft, 2026 Work Trend Index, 5 May 2026: 20,000 AI-using knowledge workers in 10 countries, Feb to Apr 2026. The report calls its links “statistical associations, not causal effects”.

26%

Only 26% of AI users say their leadership is “clearly and consistently aligned on AI”.

The same survey: Microsoft, 2026 Work Trend Index.

17

In a Microsoft-led study of 1,800 workers (Jul 2025), when managers actively modelled AI use, employees reported a 17-point lift in the AI value they reported.

Reported in the 2026 Work Trend Index. The scores are self-reported.

3/4

Nearly three-quarters of McKinsey's “AI high performers”, the 6% of organisations that credit AI with at least 5% of EBIT, have fundamentally redesigned workflows. One in five respondents say AI costs, token costs included, already constrain their use.

McKinsey survey of 1,719 people, reported by Fortune, 2 Sep 2026.

40%

A forecast that over 40% of agentic AI projects will be cancelled by the end of 2027, “due to escalating costs, unclear business value or inadequate risk controls”, with a warning about “agent washing”: products relabelled as agents without real agent abilities.

Gartner, press release of 25 Jun 2025. A forecast, not a count.

Note

Read these with care. Microsoft and McKinsey sell AI services, and their surveys are self-reported. Gartner's figure is a forecast.

Taken together they point one way, and it is why this plan starts with decisions and rules, not with more licences.

WHAT THIS MEANS FOR YOU

If the tools are bought and little has changed, the next step is not another tool: it is three decisions, a few rules and a way to measure.

Three jobs only leaders can do

IT can switch the tools on, and keen staff will find uses. Neither can decide what AI is for, set rules everyone keeps, or show by example that it is safe to try. Those are yours.

01

Decide

What AI is for here, in one sentence. Which platform comes first (page 7). Which three jobs go first, and why those. Who owns the rollout, with real time in their week. What you will do with the time it saves.

DAYS 1 TO 30

02

Set the rules

What never goes into an AI tool. Who checks AI work before it leaves the team. What AI may do on its own, and what it must ask first. Who to tell when something goes wrong.

DAYS 1 TO 30, THEN MONTHLY

03

Use it yourself

One real task of your own each week, briefed and checked, shown to your team, including what went wrong. People copy what their manager does, not what the policy says.

EVERY WEEK

What your people learn

Everyone gets the same three habits, in the same words, from the first week:

BRIEF IT

CHECK IT

KEEP IT SAFE

The 2026 AI Starter (free) teaches them in one short read, and the two workshops in days 31 to 60 practise them with real work.

Why it is a leader's job now

As of Oct 2026, Claude, ChatGPT and Gemini all take a goal and do the work across your files and apps: Claude in any conversation since 16 Sep 2026, ChatGPT in its Work mode, Gemini in its tasks. That makes the brief, the limits and the final check management questions.

An agent isn't a smarter AI. It's the same kind of AI, on a longer leash. And holding that leash is your job.

Neosaipien, checked against the UK NCSC's guidance on agentic AI, 20 Aug 2026.

THIS WEEK, IF YOU DO NOTHING ELSE

01 **Name the owner** and tell them they have 2 to 4 hours a week for 90 days.

03 **Send the questions to IT** on page 25, and ask for answers in writing by week 4.

02 **Book 90 minutes** with the leadership team for next week. Send the deck and the worksheet.

04 **Use it once yourself** on a real task. Brief it, check it, and tell your team what the check found.

The 90 days at a glance

Three stages and a monthly rhythm. Each step has a page in this plan, and a line on the 90-day checklist.

DAYS 1 TO 30	DAYS 31 TO 60	DAYS 61 TO 90	AFTER DAY 90
Decide and set the rules	Pilot	Scale, fix or stop	The monthly rhythm
01 Name an owner 02 Hold a leadership alignment session 03 Fill in the tools and connections register 04 Adopt the policy 05 Choose three first jobs with the job picker 06 Take baseline measures 07 Put the questions to IT or the admin	08 Start with two teams 09 Run the two workshops 10 Levels 1 and 2 only 11 Hold a weekly 20-minute review 12 Decide what to do with the time saved 13 Managers model it	14 Make a decision for each job 15 Write the team's way of doing a task as a skill 16 Move up one level after five clean runs 17 Keep the record of AI training 18 Report to the board	19 A monthly leadership update 20 Recheck the register and the settings 21 One job in, one job decided 22 Review the policy and the training record

WEEK BY WEEK

	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13
Name the owner, hold the session													
Questions to IT, answers by week 4													
Tools and connections register													
Policy and staff guide													
Job picker: three first jobs													
Baseline measures													
Both workshops, before any run													
Pilot jobs at levels 1 and 2													
Weekly 20-minute review													
Skills for the jobs that work													
A decision for each job													
Report to the board													

Weeks 1 to 4 are days 1 to 30, weeks 5 to 8 days 31 to 60, weeks 9 to 13 days 61 to 90. In the pilot: levels 1 Read and 2 Draft only, no job that decides something about a person, “Ask first” set in the tool, and first runs in office hours.

DAY 30 An owner, ten agreed answers, a main platform, a register, a policy, three jobs with a baseline, IT's answers.	DAY 60 Two trained teams, eight weekly reviews, a log of every surprise, each job's brief and check.	DAY 90 A decision and numbers for each job, skills for the ones that work, the training record, a report to the board.
---	--	--

Claude, ChatGPT or Gemini: what each is best at

As of Oct 2026, all three take a goal and do the work, and each is strongest somewhere. This is the honest comparison; my recommendation is on the next page. **For company data, use the tool your organisation approves.**

Task	Claude is best at	ChatGPT is best at	Gemini is best at	For most office workers
Writing	Clear business writing (Anthropic's claim for Opus 5.5); Claude Docs on every plan	Fast drafts in Chat; ChatGPT for Word even on Free	Writing inside Gmail and Docs at Google companies	Claude
Documents	Finished Word, PDF or Google Docs from one conversation; Claude for Word	Reports and Pages from Work	Gemini in Docs for Workspace users	Claude
Spreadsheets	Claude for Excel: cell references, formula-safe edits, overwrite warnings	Excel and Google Sheets add-ins; the Data plugin	Gemini in Sheets for Workspace users	Claude
Slides	Claude Slides on every plan; Claude for PowerPoint	Following company templates (OpenAI's claim for Astra); PowerPoint add-in	Gemini in Slides	Claude, with ChatGPT close
Research	Research that mixes the web with your connected work tools	Most control: edit the plan, limit it to named sites	Free Deep Research; Gemini Notebook for your own sources	ChatGPT for control over sources; Gemini Notebook for your own files
Routines and software	Claude Code; scheduled tasks and Dispatch for office routines	Codex; tasks that start on a new email, Slack message or pull request	Antigravity; schedules in tasks (personal accounts)	Claude for building; ChatGPT for routines that start on an event
Meetings	Working from a transcript you give it	Meetings plugin (beta, Mac) and audio uploads	Take notes for me in Google Meet	Gemini if you meet in Google Meet, otherwise ChatGPT
Email	Gmail connector; Claude for Outlook	Gmail and Outlook Email apps; tasks that start on a new Gmail message	Gemini built into Gmail	Gemini inside Gmail; Claude or ChatGPT for Outlook
Images	Designs, mockups and diagrams (Claude Design)	ChatGPT Images 2.5	Nano Banana images; Gemini Omni video	ChatGPT or Gemini

WHERE THE WORK HAPPENS IN EACH

CLAUDE

One conversation for questions and handed-over tasks since 16 Sep 2026, rolling out with Pro and Max first. Code, a separate place, for building software only.

Each company publishes tests it wins. Anthropic reports Opus 5.5 at 66.4% on Terminal-Bench 4.0, OpenAI reports GPT-6 Astra at 57.9% on the same test, and Google says Argon ranks first on the Vals Index. Do not choose on vendor numbers alone. Claims marked as a vendor's are that vendor's own.

CHATGPT

One app since 9 Jul 2026: Chat for quick help, Work for anything that ends in a file, a report or a done task, Codex for software.

GEMINI

At work, inside Gmail, Docs, Sheets, Slides, Drive, Chat and Meet. Its Task agent is for personal accounts on Google AI Pro or Ultra, for now. Antigravity for software.

MY RECOMMENDATION

Claude first, ChatGPT second, Gemini third.

My own view for most organisations, as of Oct 2026, next to the comparison on page 6. The tool you already approve for company data still comes first.

WHY CLAUDE FIRST

- The simplest start to finish: one conversation for quick questions and handed-over tasks, with no mode to choose.
- It asks before acting by default, and Anthropic reports Opus 5.5 is harder to steer with hidden instructions and less likely to take hard-to-reverse actions.
- Finished office files on every plan, Free included: Docs, Slides and Design. Claude for Excel cites cells and keeps formulas intact.

WHY CHATGPT SECOND

- ChatGPT Work is a strong agent: a cloud browser for signed-in websites, and scheduled tasks that start when an email or message arrives.
- A wide set of everyday features: voice, images, Sites, and ChatGPT for Word even on Free.
- OpenAI's own tests put GPT-6 Astra ahead on computer use, and OpenAI calls it its best model for following company templates.

WHY GEMINI THIRD

- The best fit when a company already runs on Google Workspace: Gemini in Gmail, Docs, Sheets and Meet.
- A generous free plan with Deep Research, Gemini Live and Canvas.
- But its agent tasks, custom apps and skills are limited to personal accounts for now, and many features start in the US and in English only.

Roll it out in that order

Days 1 to 30. One main platform for the pilot. Claude, unless the organisation has already standardised on another tool. Put it in the register, choose one plan for the pilot teams, and send IT the questions on page 25.

Days 31 to 60. Pilot on the main platform only. Run both workshops in it. Set “Ask first” and the connector approvals in it. Levels 1 and 2 only.

Days 61 to 90. ChatGPT second, for named jobs. Only where it wins: routines that start when an email or Slack message arrives, research where you must control the sources, pictures. Each one goes through the job picker, with its own baseline.

After day 90. Gemini third, where you run Google Workspace. Notes in Google Meet and writing inside Gmail. Its Task agent is for personal accounts only, as of Oct 2026, so keep agent jobs on the main platform.

WHEN TO CHANGE THE ORDER

- Already standardised on one tool: use that tool, and run this plan inside it.
- On Google Workspace: use Gemini inside Gmail, Docs and Meet from day 1 for those jobs. Claude also works in Google Docs, Sheets and Slides (Claude for Google Workspace, public beta, paid plans).
- Two platforms at once doubles the register, the questions for IT and the training. Add the second only when a named job needs it.

Five things to set up for everyone

Whichever platform you choose, staff meet these in their first week. Decide them once, centrally, in plain words, and write the answers in the policy. All as of Oct 2026.

01 CONNECTORS AND MCP

A connector is a signed-in link from the AI to another app: files, calendar, email. It reaches only what that person's account can reach, and only what the connector allows. Claude calls it a connector, ChatGPT an app inside a plugin, Gemini a connected app. They use one shared standard, MCP, which its own site compares to a USB-C port for AI applications.

Decide: which connectors are allowed, read-only first, and who adds them. On Claude Team and Enterprise an owner adds connectors first; on a Google work account the admin decides which apps are allowed.

02 PROJECTS

A project holds one job's files, instructions and chats, so nobody pastes the same background twice. Claude and ChatGPT call it a project and can share one with view or edit access; Gemini's nearest is a notebook, shared with Gemini Notebook.

Decide: one shared project for each pilot job, owned by its manager.

03 SKILLS

A skill is a saved how-to: one SKILL.md file that all three accept. The AI reads its short description, and loads the rest only when a request matches.

Decide: the team's way of doing a job becomes a skill in days 61 to 90 (step 15, page 19).

04 MEMORY AND INSTRUCTIONS

Each person can save standing instructions, and the AI can remember past chats. On Claude Team and Enterprise an owner switches memory on and each person opts in; on a Google work account the admin controls it and keeps chat history 18 months by default; in ChatGPT it sits under Settings, Personalization, Memory.

Decide: memory on or off for company work.

05 WHICH MODEL WHEN

The job	Claude	ChatGPT	Gemini
Quick question, short rewrite, sort a list	Haiku 5.5 or Sonnet 5.5	GPT-6 at Instant	Auto or Flash-Lite
Everyday writing, a summary, a first deck	Sonnet 5.5 at its Default effort	GPT-6 at Medium	Flash
A careful report, a messy spreadsheet	Opus 5.5	GPT-6 at High	Pro, thinking High
Hand over a multi-step task	Opus 5.5, in an ordinary conversation	Work with Default, or GPT-6.1 Sol	Task with Flash or Argon
The hardest, high-stakes work	Fable 5.1 (it uses limits fastest)	GPT-6 at Pro, or GPT-6 Astra in Work	Argon, thinking High

Decide: start everyone on the default and move up only when the answer misses something that matters. Heavier models and task work use limits faster, and not every model is on every plan: Opus 5.5 needs a paid Claude plan, Fable 5.1 needs Max, Team Premium or Enterprise, GPT-6 at Pro needs Pro, Business or Enterprise, and Argon needs Google AI Pro or Ultra. For staff, the handout in Workshop in a Box: Brief an Agent shows each step in each tool.

Decide and set the rules

The first month changes nothing in anyone's daily work. It decides what AI is for, puts the rules in writing and measures today's work, so that day 90 has something to compare with.

- | | |
|--|--|
| 01 Name an owner | 02 Hold a leadership alignment session |
| 03 Fill in the tools and connections register | 04 Adopt the policy |
| 05 Choose three first jobs with the job picker | 06 Take baseline measures |
| 07 Put the questions to IT or the admin | |

BY DAY 30 YOU HAVE

- A named owner with time to do it
- Ten agreed answers from the leadership team
- A register of every AI tool in use and what it connects to
- A policy, and a one-page guide everyone has read
- Three first jobs, each with a level, an owner and a check
- A two-week baseline for each job, and IT's answers in writing

THE TRAP TO AVOID

Starting the pilot in week 1 because people are keen. Without a baseline you cannot tell at day 90 whether anything improved, and without rules the keen few set them by habit.

Seven steps. Every time on these pages is an estimate.

01 Name an owner

WHO **The leader** WHEN **Days 1 to 3** TIME (ESTIMATE) **30 minutes to decide**

Pick one person, senior enough to say no to a team that wants to skip a step, with about 2 to 4 hours a week for the 90 days. Often an operations lead. Not IT on its own: the questions are about work, not software.

THE OWNER

- keeps the register and the record of AI training
- chairs the weekly reviews with the pilot managers
- writes the monthly update
- can pause any job at once

THE OWNER DOES NOT

- approve every use of AI by everyone
- choose the jobs alone
- own the rules alone: the leadership team does

02 Hold a leadership alignment session

WHO The leadership team WHEN Week 1 TIME (ESTIMATE) 90 minutes

One meeting of the leadership team, in week 1, to agree ten answers. Use the kickoff deck to open it and the worksheet on the next page to close it. Agree each answer in the room, or name who decides it and by when.

MIN	What happens	Use
10	Why now, and why it is the leaders' job	Deck slides 1, 2 and 4
15	Where we are today: tools, users, rules, measures	Deck slide 3, the register so far
40	The ten questions, one answer each	The worksheet
15	The main platform, the 90 days and the first three jobs	Deck slides 5 to 7
10	Decisions, owners and dates; each leader names one task of their own	Deck slides 8 and 9

About 90 minutes in all. Send the deck and the worksheet two days before, so people arrive with views.

USE The worksheet and the kickoff deck in this pack

RUNNING IT WELL

- Start from the register: what people already use is the honest starting point.
- One answer per question. A list of everyone's views is not an answer.
- The most senior person speaks last on each question.
- Write the answers on the worksheet as you go, in the words the room agreed.
- End with each leader naming one task of their own for this week.

IF THE ROOM CANNOT AGREE

- On what AI is for: take the narrower answer for 90 days and revisit at day 90.
- On the first jobs: take the lower-risk job; the picker's score breaks ties.
- On what never goes in: the stricter rule wins until the policy is adopted.
- On anything else: name who decides, and by which date, before you leave.

WHAT GOOD LOOKS LIKE

Every leader can say, in the same words, what AI is for here, which three jobs go first, and what never goes in. Ask them a week later.

Leadership alignment worksheet

Fill it in together in the session. One answer per question, agreed by the room, not a list of everyone's views. Also in this pack as its own page to print.

01 What is AI for here, in one sentence?

For example: "Faster, better answers for customers, and time back for [the work we never get to]."

03 Which three jobs first, and why those?

From the job picker shortlist. Two teams at most.

05 What never goes into an AI tool here?

The top three, in plain words. The policy holds the full list.

07 What will we do with the time saved?

Decide before it is saved, team by team.

09 What will each of us use it for ourselves this month?

One real task each, briefed and checked, shown at a review.

02 Which AI platform is our main one, and which comes second?

The plan's author recommends Claude first, ChatGPT second, Gemini third (the plan, page 7). The tool we already approve for company data comes first.

04 Who owns the rollout, and how much of their week?

A named person, not a committee. About 2 to 4 hours a week (an estimate).

06 What is the most AI may do in the next 90 days?

This plan says level 1 Read and level 2 Draft only in the pilot.

08 How will we know at day 90 that it worked?

Quality, rework, turnaround, errors caught, staff confidence. Against a baseline.

10 What would make us stop a job?

For example: worse quality after a fix, or a risk we can't manage.

Agreed on [date]

Owner [name]

Next review [date]

03 Fill in the tools and connections register

WHO The owner, with IT or the admin WHEN Weeks 1 and 2 TIME (ESTIMATE) 2 to 4 hours

A list of every AI tool anyone uses for work, paid for or not, and what each one can reach. It is the first thing IT and the leadership team will both want, and it usually turns up tools nobody approved.

ONE ROW PER TOOL

Tool and plan	Who may use it	Data allowed	Connected apps	Can they write?
Approval setting	Memory	Scheduled tasks	Owner	Review date

Tool and plan	Who	Data allowed	Connected apps	Write?	Approval	Memory	Scheduled	Owner
[Assistant], business plan	All staff	Internal, no personal data	Shared drive	Read only	Ask first	Off	None	[owner]
Note-taker inside [video app]	Found in use by 4 people	Not yet decided	Calendar	Yes	Not known	Not known	Every meeting	None yet

EXAMPLE, MADE UP The second row is the kind of thing a register finds. Ask each team “what do you use, even once a week?” rather than “what is approved?”. Record what is true today, not what should be true. Look in the places tools hide: browser extensions, meeting note-takers, AI features switched on inside software you already pay for, and personal accounts used for work.

USE The Team AI Policy Kit's register, or the short register in this pack's workbook

04 Adopt the policy

WHO The leadership team WHEN Weeks 2 and 3 TIME (ESTIMATE) 2 hours to adapt, 1 meeting to agree

A short written policy, agreed by the leadership team, and a one-page guide that every person reads. At the least it should say:

- which tools are approved, and how to ask for a new one
- what never goes in, in plain words, with examples
- who checks AI work before it leaves the team
- when you tell people that AI made something
- what agents may do: “Ask first” is set in the tool, not typed in the brief
- outside words, no power: an agent that reads outside content may not also send, pay or delete without asking
- an agent found sending, paying or deleting without asking is set to “Ask first” the same day
- who to tell when something goes wrong, with no blame for telling quickly
- that AI never makes a decision about a person on its own

The last line is this plan's recommendation, not a quote from any law. Send everyone The 2026 AI Starter (free) with the guide, so the three habits arrive with the rules.

USE The Team AI Policy Kit; The 2026 AI Starter for everyone

05 Choose three first jobs with the job picker

WHO The owner and two team managers WHEN Weeks 2 and 3

TIME (ESTIMATE) 60 minutes per team

Each pilot team lists 10 to 15 tasks it does, then scores each one in the job picker. It takes about an hour per team. The workbook works the scores and the ceiling out for you.

USE The job picker in this pack

1. VALUE, 1 TO 5

- 1 Little difference to anyone
- 2 Saves a person some effort
- 3 Noticeably faster or better for the team
- 4 Noticeably better for customers or for money
- 5 Changes how the team works

2. HOW OFTEN, 1 TO 5

- 1 A few times a year
- 2 Monthly
- 3 Weekly
- 4 Daily
- 5 Many times a day

3. THE THREE QUESTIONS SET THE CEILING

01 Can it be undone easily?

If not: level 3 at most.

02 Does it leave the team?

A customer, a supplier, the public or money. If so: level 3 at most, and a person presses send.

03 Does it read outside content?

The web, outside email, files others sent. If so: it may not also send, pay or delete without asking.

Take the lowest level any of them allows.

4. WHICH TOOL

CHAT

One answer, draft or idea that you use yourself.

DEEP RESEARCH

A cited report on a question, where checking the sources is the work.

AGENT

Work whose next step depends on what it finds along the way.

AUTOMATION

The same fixed steps every time, written down in advance.

Use the smallest tool that does the job. Gartner's advice (25 Jun 2025): "using AI agents when decisions are needed, automation for routine workflows and assistants for simple retrieval".

The rule. Score = value x how often. Shortlist the highest scores, then take the lowest level the three questions allow. Pick three jobs across two teams, each at level 1 or 2 for the pilot.

Never in the first 90 days. Leave out of the first 90 days any job that decides something about a person: hiring, promotion, pay, discipline, task allocation or performance. Those carry their own rules (see Rules for leaders). Not legal advice.

THE SCORING HOUR

10 MIN

List the tasks the team does every week, in its own words.

20 MIN

Score value and how often. Argue about the 4s and 5s, not the 1s.

15 MIN

Answer the three questions for the top eight.

15 MIN

Pick the tool for each, then the two or three for the pilot.

EXAMPLE, MADE UP

The job picker, filled in

A made-up services firm of about 60 people. It names no real organisation. Two teams scored eight tasks, in about an hour each.

Q1 Can it be undone easily? Q2 Does it leave the team? Q3 Does it read outside content?

	Task and team	Value	Often	Score	Q1	Q2	Q3	Ceiling	Tool	Pilot	Pick
1	First drafts of replies to routine customer questions Customer service	4	5	20	No	Yes	Yes	3 Ask first	Agent	2 Draft	Yes
2	Proposal first drafts from our template and past proposals Sales	5	3	15	No	Yes	No	3 Ask first	Chat	2 Draft	Yes
3	Weekly summary of themes in customer feedback Customer service	3	3	9	Yes	No	Yes	3 Ask first	Agent	1 Read	Yes
4	Monthly management report from three systems Finance	4	2	8	Yes	No	No	4 Alone	Automation		Later
5	Check expense claims against the policy Finance	3	3	9	Yes	No	No	4 Alone	Automation		Later
6	Weekly look at competitors' public pages Marketing	2	3	6	Yes	No	Yes	3 Ask first	Agent		No
7	Research a new supplier market and shortlist Operations	3	1	3	Yes	No	Yes	3 Ask first	Deep research		No
8	Screen job applications HR	4	3	12	No	Yes	Yes				Never

Why these three

- **First drafts of replies to routine customer questions.** Highest score. Drafts only: a person reads and sends every reply.
- **Proposal first drafts from our template and past proposals.** High value. The sales lead checks the facts and the figures before anything leaves.
- **Weekly summary of themes in customer feedback.** Reads and reports only. Uses the same team as job 1, so one review covers both.

Why not the others, yet

- **Monthly management report from three systems.** The steps are the same every month, so it is an automation project, not a pilot job.
- **Check expense claims against the policy.** Fixed checks suit automation; a person decides every exception.
- **Screen job applications.** It decides something about a person. See Rules for leaders.

All three picks have a ceiling of level 3: jobs 1 and 2 leave the team and cannot be taken back once sent, and jobs 1 and 3 read customers' emails. In the pilot they stay at 1 or 2. The workbook holds the same rows, marked as an example.

06 Take baseline measures

WHO The two team managers WHEN Weeks 3 and 4 TIME (ESTIMATE) 2 weeks of normal work

Before anyone uses AI on the three jobs, measure how the work goes today. Two weeks of normal work is usually enough. Use the same method you will use at day 90, or the comparison means nothing.

QUALITY How good the finished work is.	REWORK How often the work has to be done again.	TURNAROUND How long from request to finished work.	ERRORS CAUGHT How many mistakes the checks stop before work goes out.	STAFF CONFIDENCE Whether people know when to trust the work and when not.
--	---	--	---	---

IN PRACTICE

- One person per team scores quality, the same person at day 90
- Take samples from normal weeks, not the busiest
- Write each rubric down before the first score

KEEP IT SMALL

- Twenty pieces of work a week per job is plenty
- Turnaround from timestamps you already have
- Staff confidence: one question, the same each time

How to measure each one, with examples, is on page 22. Note time saved as well, but do not treat it as the result.

USE The measures tracker in this pack

07 Put the questions to IT or the admin

WHO The owner WHEN Send in week 1, answers by week 4 TIME (ESTIMATE) 1 meeting

Send the questions on page 25 in week 1 and ask for answers in writing by week 4. They cover who owns agents, how agents are identified and logged, which connections are allowed and whether they can write, memory, where data goes and for how long, and how to stop everything at once.

Where an answer is “we don't know”, the pilot stays at levels 1 and 2 for that tool until it is known, and the owner notes it in the register. That is not a reason to delay the plan; it is the plan working.

USE Questions for IT or your admin, in this pack

READY FOR DAY 31 WHEN

Every line of days 1 to 30 on the checklist is ticked, or the leadership team has agreed in writing why one is not.

Pilot

Two teams do the three jobs for real, at levels 1 and 2 only, and look at the results together every week. The point is to learn what works here, not to prove that AI works.

08 Start with two teams

10 Levels 1 and 2 only

12 Decide what to do with the time saved

09 Run the two workshops

11 Hold a weekly 20-minute review

13 Managers model it

08 Start with two teams

WHO The owner WHEN Week 5

Two teams, so you can compare: one keen, one typical, each with a manager willing to use AI in the open and to stop a job that is not working. Each job gets a six-part brief: Goal / Done looks like / Context / May touch, must not / Limits / Stop and ask.

09 Run the two workshops

WHO Each pilot team WHEN Week 5, before the first run TIME (ESTIMATE) 60 minutes each

Week 5, before the first run: Brief an Agent, then Check What AI Writes, sixty minutes each, from the boxes, or I can run them live (page 27). **Without the boxes,** spend each hour on one real task, briefed in the six parts above and checked with the five checks on page 17. Log each in the record of AI training.

USE Workshop in a Box: Brief an Agent; Workshop in a Box: Check What AI Writes

10 Levels 1 and 2 only

WHO Each pilot team WHEN Weeks 5 to 8

■■■■

1 READ

It looks and reports. It changes nothing.

■■■■

2 DRAFT

It drafts. Only a person sends or shares.

■■■■

3 ASK FIRST

It acts, and asks before each send, pay, delete or change.

■■■■

4 ALONE

It acts inside fixed limits. A person reviews after.

“Ask first” set in the tool; nothing scheduled that sends; first runs in office hours. The UK’s cyber agency suggests “running initial experiments in office hours, when more human oversight is available” (NCSC, 20 Aug 2026).

USE The Agent Playbook for each pilot manager, or the brief above and these two levels on their own

WHAT COUNTS AS A SURPRISE

It touched something the brief did not name, sent or tried to send, invented a fact or a quote, got a figure wrong against its source, or carried on where it should have stopped and asked.

WHAT TO DO AT ONCE

Stop the job, look at where it acted, tell the owner the same day, drop it one level and log one line. Nobody is blamed for reporting quickly.

11 Hold a weekly 20-minute review

WHO Each pilot team and the owner WHEN Every week TIME (ESTIMATE) 20 minutes

Same time each week, each pilot team with its manager, the owner joining. It is short on purpose.

MIN	What happens
5	The numbers. Runs this week, clean runs, surprises. A run is one go at the job, from the brief to the checked result; a clean run passed the five checks with no correction that mattered.
10	One real example. Someone shows a brief, what came back, and what their check found. Pick a bad one as often as a good one.
5	Decisions. Change the brief, the context or the check; pause a job; note what to raise with the owner.

THE FIVE CHECKS

1 The job. Did it do what the brief asked?	2 The facts. Is every name, number, date and quote real?	3 The steps. How did it get there?	4 The changes. What is different now?	5 The sign-off. Would you put your name to it?
--	--	--	---	--

THE REVIEW LOG, ONE LINE A WEEK

Week	Job	Runs	Clean runs	Surprises	Example shown	What we changed
Raised with the owner						

USE The AI Fact-Check Kit for each pilot manager

12 Decide what to do with the time saved

WHO The leadership team

WHEN Before week 5

TIME (ESTIMATE) 15 minutes per team

Decide before the pilot, team by team, where the time goes. Otherwise it disappears.

- Back into the same work: faster replies, more customers seen
- To a named job the team never gets to
- Off the backlog or the overtime

13 Managers model it

WHO Every leader and pilot manager

WHEN Every week

TIME (ESTIMATE) One real task a week

Each leader and pilot manager uses AI on one real task of their own each week, and shows it at a review: the brief, the check, and what went wrong.

- Ask “what was checked?” whenever AI work reaches you
- Never accept “The AI did it” as the answer to “who checked this?”

Scale, fix or stop

Each job is decided on its numbers, against its baseline. Stopping a job that does not work is a result, and it is cheaper now than in a year.

14 Make a decision for each job

16 Move up one level after five clean runs

18 Report to the board

15 Write the team's way of doing a task as a skill

17 Keep the record of AI training

14 Make a decision for each job

WHO The leadership team WHEN Days 75 to 80 TIME (ESTIMATE) 45 minutes

One meeting, around days 75 to 80, with each pilot manager bringing the measures tracker.

SCALE

Quality and rework are at least as good as the baseline, turnaround or errors caught are better, there is no open surprise, and the team wants to keep it.

Then: write it as a skill, offer it to a second team, keep measuring.

FIX

Promising, but one measure is worse, or the checks keep finding the same kind of mistake.

Then: change one thing (the brief, the context, the check or the level) and run it three more weeks.

STOP

No gain after a fix, a risk you cannot manage at a level you accept, or the team does not trust it.

Then: write down why in two lines, and take the next job from the picker.

SIX QUESTIONS FOR THE MEETING

01 Did quality hold, on the same rubric, scored by the same person?

03 Where did the time saved go, and was that what we decided?

05 Would the team miss it if we stopped tomorrow?

02 Did rework fall, or move somewhere else?

04 What did the checks catch, and is it the same mistake each week?

06 What would have to be true to move it up one level?

USE The measures tracker in this pack

15 Write the team's way of doing a task as a skill

WHO The pilot manager WHEN Days 61 to 85 TIME (ESTIMATE) 2 to 3 hours per skill

A skill is a saved, reusable set of instructions the assistant follows each time it does one kind of task. It replaces the list of prompts people used to pass round. As of Oct 2026, Claude, ChatGPT and Gemini all accept the same format, a SKILL.md file, first developed by Anthropic and released as an open standard: Claude under Customize, Skills; ChatGPT inside plugins, where typing @skill-creator starts one; Gemini under Customize, Skills, on personal accounts for now.

WHAT GOES IN

- When to use it, and when not to
- The six-part brief: Goal / Done looks like / Context / May touch, must not / Limits / Stop and ask
- One good example of the finished work
- The check a person runs before it goes out
- The level, the owner and the date it was last reviewed

EXAMPLE, MADE UP

Draft a reply to a routine customer question

Goal: a reply the agent drafts and a person sends, within [4] working hours.

Context: the reply rubric, the [help pages], three good past replies.

May touch, must not: read the support inbox; never send, refund or promise a date.

Stop and ask: a complaint, a refund, anything about a person's account.

Level: 2 Draft. **Owner:** [customer service manager].

The 20 briefs in The Agent Playbook come as skill files to start from, and its dated insert covers each tool's settings.

USE The Agent Playbook's briefs as skills

16 Move up one level after five clean runs

WHO The pilot manager WHEN From day 61

■■■■

1 READ

It looks and reports. It changes nothing.

■■■■

2 DRAFT

It drafts. Only a person sends or shares.

■■■■

3 ASK FIRST

It acts, and asks before each send, pay, delete or change.

■■■■

4 ALONE

It acts inside fixed limits. A person reviews after.

A **run** is one go at the job, from the brief to the checked result. A **clean run** is one where the five checks found nothing that mattered. After five in a row, a job may move up one level, never above the ceiling its three questions set. After any surprise it goes down a level at once, and the next five runs start again. Level 3 only with "Ask first" set as an approval in the tool.

17 Keep the record of AI training

WHO The owner WHEN From day 31, checked at day 85 TIME (ESTIMATE) 10 minutes a session

One line for every session anyone attends: date, who, what, the materials, how long, and who led it. Start it on day 31 with the workshops, and check it is complete by day 85.

Date	Who	What	Materials	Length	Led by
[date]	Customer service team, 9 people	Check What AI Writes	Workshop in a Box	60 min	[team manager]
[date]	Sales team, 6 people	Brief an Agent	Workshop in a Box	60 min	[team manager]

EXAMPLE, MADE UP Two rows of a made-up record.

THE EU RULE BEHIND IT

Since July 2026, Article 4 of the EU AI Act, as replaced by Regulation (EU) 2026/1744, says providers and deployers of AI systems “shall take measures to support the development of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf”. It adds: “This obligation does not require providers or deployers to guarantee any specific level of AI literacy of any individual.”

A record shows the measures you took. Whether it applies to you is a question for your adviser. Not legal advice.

USE The Team AI Policy Kit's record of AI training

18 Report to the board

WHO The leader and the owner WHEN Day 90 TIME (ESTIMATE) One page, 20 minutes to present

One page, from the measures tracker, presented in about 20 minutes. Use the monthly update template's headings.

- 01 What we set out to do. The one sentence from the worksheet, and the three jobs.
- 02 What changed. Each job's five measures against its baseline, in a small table.
- 03 What we decided. Scale, fix or stop for each job, with the reason.
- 04 What went wrong. Every surprise, what we did at once, what we changed.
- 05 The rules. Policy adopted, register current, training recorded.
- 06 The next 90 days. The next jobs, the teams, and what we need from the board.

WHAT THE BOARD WILL ASK, SO ANSWER IT FIRST

- What did it cost, in licences and in people's time?
- What went wrong, and how do we know we would see it?
- What did we stop, and why?
- Who owns this now, and what happens in the next 90 days?

USE The monthly update template in this pack

The monthly rhythm

After day 90 the plan becomes a habit: one short meeting a month, one job in at a time, and the rules kept as current as the tools.

19 A monthly leadership update

21 One job in, one job decided

20 Recheck the register and the settings

22 Review the policy and the training record

19

A monthly leadership update

Monthly, about 30 minutes, run by the owner from the one-page template: the jobs, the measures, surprises, rules and tools, training, decisions needed.

20

Recheck the register and the settings

Monthly, with IT or the admin. Check what each tool connects to, its approval and memory settings, and what the vendor changed or retired. Update the register.

21

One job in, one job decided

When a job is scaled or stopped, the next one comes from the picker, at level 1 or 2, with its own baseline. One at a time keeps the review short.

22

Review the policy and the training record

Every three months: does the policy still match what people do? Is every new starter in the training record? A short report to the board.

SIGNS IT IS WORKING

- People bring AI work to the review without being asked
- Quality holds while the same mistakes stop coming back
- New tools arrive through the register, not by surprise
- Every leader can name the jobs and their numbers
- Nobody says "The AI did it"

KEEPING UP WITH THE TOOLS

Every vendor fact here is dated: as of Oct 2026. Follow the changes in the vendors' own release notes (page 26).

DAYS 91 TO 180

- Offer each job that scaled to a second team, with its own baseline
- Bring in the next jobs from the picker, one at a time
- Move a job up only after five clean runs, never above its ceiling
- Report to the board again at day 180

Time saved is not value

Hours saved is the easiest number to collect and the easiest to fool yourself with. A faster draft that needs rework, or time saved that nobody uses, changes nothing. Measure the work, with a baseline first.

Measure	What it tells you	How to measure it	EXAMPLE, MADE UP
Quality	How good the finished work is.	Score a sample against one short rubric of three to five points, the same rubric before and after.	Score 20 replies a week from 1 to 5 against the team's reply rubric.
Rework	How often the work has to be done again.	Count the pieces that had to be redone or corrected after they left the team.	Count replies that needed a second reply because the first was wrong.
Turnaround	How long from request to finished work.	Take the median time for the same kind of task.	Median working hours from a customer's question to our reply.
Errors caught	How many mistakes the checks stop before work goes out.	Log each error found at the check, with what it was.	Wrong dates, figures quoted from memory, names misspelt, found before sending.
Staff confidence	Whether people know when to trust the work and when not.	Ask one question, the same each time, scored 1 to 5.	"I know what to check before I send an AI draft."

FOUR RULES FOR THE NUMBERS

- **Baseline first:** two weeks before the pilot, by the same method.
- **Same sample, same rubric:** the person who scored the baseline scores day 90.
- **Write down the misses:** every surprise counts, even one with no harm done.
- **Time saved goes in the notes,** with where the time went.

WHY THIS MATTERS NOW

A McKinsey survey found the firms getting the most from AI far more likely to redesign workflows rather than add AI to existing processes. Time saved becomes value only when someone decides what it is for, and the numbers show whether it did.

McKinsey, reported by Fortune, 2 Sep 2026. McKinsey sells AI services, and the survey is self-reported.

ONE JOB AT DAY 90

EXAMPLE, MADE UP

Customer reply drafts	Baseline	Day 90	Reads as
Quality, rubric 1 to 5	3.6	3.8	Slightly better
Rework: second replies needed, per 100	7	5	Better
Turnaround, median working hours	6.5	3	Better
Errors caught at the check, per 100	2	6	The check is working; look at what it catches
Staff confidence, 1 to 5	2.9	3.7	Better

Made-up numbers, to show how to read them. The decision would be scale, with a look at why drafts carry errors that the check now catches.

What it costs, beyond the licences

Licences are the visible cost. The others are usage limits, the time to brief and check, and agent projects that never pay back. Plan for all three from the start.

THE DIAL ON THINKING

Some assistants now let people choose how hard the model thinks. Anthropic's help page for Claude, as of Oct 2026: "Higher effort means more thorough responses, but they take longer and use more tokens, so you'll reach your usage limits faster." And: "Low and Medium work well for routine tasks and stretch your usage further."

The same dial, as of Oct 2026: Effort in Claude, the Thinking slider in ChatGPT, a thinking level in Gemini.

Anthropic, Change the model, effort, and thinking settings; OpenAI and Google model help pages; all read 16 Oct 2026.

COST IS NOW A LIMIT

One in five respondents to a McKinsey survey say AI-related operating costs, including token costs, are already constraining their use. Gartner's 2025 forecast that over 40% of agentic AI projects will be cancelled by the end of 2027 names "escalating costs" first among its reasons.

McKinsey survey of 1,719 people, as reported by Fortune, 2 Sep 2026. Gartner, 25 Jun 2025 (a forecast).

Six habits that keep cost in proportion

- 01 **Smallest tool first.** Chat before deep research, deep research before an agent, and fixed automation for fixed steps.
- 02 **Match the effort to the job.** A low setting for routine drafting; high only for analysis and planning. Put the setting in the skill.
- 03 **Count the checking time.** A draft that takes longer to check than to write is not a saving. Record both in the pilot.
- 04 **Licences follow use.** Each month, compare who has a licence with who used it. Move idle ones.
- 05 **No agent without a decision in it.** If every step can be written down in advance, it is automation, usually cheaper and more predictable.
- 06 **Watch for "agent washing".** Ask any vendor selling "agents" what it decides on its own, and what it asks a person first.

QUESTIONS FOR THE VENDOR OR YOUR RESELLER

- What are the usage limits on each of our plans, and what happens when someone reaches one?
- Do agent tasks, scheduled tasks and higher effort settings count against the same allowance?
- Can an admin see usage per person and per team, and set limits?
- What changes at renewal, and how much notice do we get?

This plan quotes no prices: they change too often and differ by plan and by country. Check your own contracts and the vendors' own pages.

Rules for leaders

NOT LEGAL ADVICE

Plain-English orientation, read from the official texts on 16 Oct 2026. Whether a rule applies to you depends on where you operate and what you use AI for. Check with your own adviser.

EU AI ACT, ARTICLE 4, AS AMENDED

Rewritten by Regulation (EU) 2026/1744 (signed 8 Jul 2026, published 24 Jul 2026). Providers and deployers “shall take measures to support the development of AI literacy” of their staff, and need not “guarantee any specific level of AI literacy of any individual”. **In this plan:** the record of AI training.

EU AI ACT, ARTICLE 50, FROM 2 AUG 2026

Deployers must disclose deep fakes, and must disclose AI-generated text “published with the purpose of informing the public on matters of public interest”, unless it “has undergone a process of human review or editorial control” and someone “holds editorial responsibility”. **In this plan:** the policy's line on when you say AI made it.

EU AI ACT, ARTICLE 5, SINCE 2 FEB 2025

Banned: AI systems “to infer emotions of a natural person in the areas of workplace and education institutions”, except for medical or safety reasons. **In this plan:** check that no tool in the register does this.

EU AI ACT, ANNEX III POINT 4, FROM 2 DEC 2027

High-risk: AI for recruitment or selection, including “to analyse and filter job applications, and to evaluate candidates”, and for decisions on promotion or termination, “to allocate tasks based on individual behaviour or personal traits or characteristics or to monitor and evaluate the performance and behaviour of persons”. Due on 2 Aug 2026, these duties now apply from 2 Dec 2027. **In this plan:** no such job in the first 90 days.

ILLINOIS, PUBLIC ACT 103-0804, SINCE 1 JAN 2026

An employer may not use AI “that has the effect of subjecting employees to discrimination on the basis of protected classes” in recruitment, hiring, promotion, discipline, discharge and similar decisions, or “use zip codes as a proxy for protected classes”, and must give notice when it uses AI for those purposes.

ASK YOUR ADVISER ABOUT

Data protection. Monitoring staff. Who owns what AI makes. Rules for your sector. Recording meetings and AI note-takers. What to tell customers.

THE PLAN'S OWN RULES

- “Ask first” is set in the tool, not typed in the brief.
- Outside words, no power: an agent that reads outside content (the web, outside email, files others sent) may not also send, pay or delete without asking.
- Before you forward it: whoever briefed or asked runs the checks, and “The AI did it” is never the answer to “who checked this?”

Questions for IT or your admin

Organisation version. Send it in week 1; ask for written answers by week 4. Where an answer is “we don't know”, the pilot stays at levels 1 and 2 until it is known.

01 WHO OWNS AGENTS

- Who may create an agent, a scheduled task or a task started by an event, and who approves a new one?
- Who is the named owner of each agent, and who takes over when that person leaves?
- Where is the list of every agent and scheduled task running for us today?

02 AGENT IDENTITIES AND THE AUDIT TRAIL

- Does each agent act under the account of the person who set it up, or under an identity of its own?
- Where is the record of what each agent read, sent and changed, and who reads it?
- How long is that record kept, and can it be changed after the fact?

03 WHICH CONNECTORS ARE ALLOWED, AND WHETHER THEY WRITE

- Which apps may our AI tools connect to, for whom, and on which plans?
- For each connection: may it only read, or also write, send or delete?
- Who approves a new connection, and can “ask first” be set as an approval in the tool?

04 THE PLATFORM AND ITS MEMORY DEFAULTS

- Which of Claude, ChatGPT and Gemini do we license, on which plans, and for whom?
- Is memory on or off by default for our accounts, and can staff change it?
- Can a person see and clear what their assistant remembers, and can we clear it for them?

05 DATA LOCATION AND RETENTION

- Where is our data processed and stored when staff use these tools, including files opened from a laptop?
- How long are chats, files, task results and logs kept, and can we set that?
- On our plans, is our data used to train the vendor's models, and where is that written?
- When a connection is removed, what happens to what the tool already took from it?

06 HOW TO PULL THE PLUG

- How do we stop one agent now, and how do we stop all of them at once?
- Who can do it outside office hours, and how long does it take?
- Can we cut every connection to one app in one step?
- When did we last try it?

Microsoft's 2026 Work Trend Index tells IT leaders to treat agents “as managed entities with identities, permissions, policy enforcement, and lifecycle management”. Microsoft, 5 May 2026

As of Oct 2026, on Claude Team and Enterprise an owner adds connectors first; on a Google work account the admin decides which apps are allowed. Anthropic says custom connectors “aren't verified”. Anthropic and Google help pages, read 16 Oct 2026

The UK's National Cyber Security Centre: “you should always be able to ‘pull the plug’ and halt autonomous AI agent activity immediately.” UK NCSC, 20 Aug 2026

Each tool's own settings and their defaults, as of Oct 2026, are in The Agent Playbook's dated insert. These products change monthly. Check the settings in your own tool.

Checked against

Checked on the vendors' and lawmakers' own pages on 16 Oct 2026. These products change monthly. Check the settings in your own tool. Each fact was read at its source; a product's behaviour is stated “as of Oct 2026”.

Source, and what was read	Address, and date read
Anthropic, on Claude. Claude Cowork and chat are now one Claude, 16 Sep 2026; connectors and the connectors directory; models, effort and plans; projects, memory and skills; Claude for Excel and for Google Workspace; Introducing Claude Opus 5.5, 22 Sep 2026.	https://claude.com/resources/articles/cowork-is-now-claude Read 16 Oct 2026
OpenAI, on ChatGPT. Release notes: ChatGPT Work, and one app for Chat, Work and Codex, 9 Jul 2026; ChatGPT Work and Codex; connected apps and their permissions; models; projects; memory; GPT-6 Astra, 3 Sep 2026.	https://help.openai.com/en/articles/20001275-chatgpt-work-and-codex Read 16 Oct 2026
Google, on Gemini. Tasks in Gemini (personal accounts); Google Workspace with Gemini; work and school accounts; connected and custom apps; skills; release notes, with Gemini 4 Argon, 13 Oct 2026.	https://support.google.com/gemini/answer/17094507 Read 16 Oct 2026
Model Context Protocol. What is the Model Context Protocol: the USB-C comparison.	https://modelcontextprotocol.io/docs/getting-started/intro Read 16 Oct 2026
Microsoft. 2026 Work Trend Index annual report, “Agents, human agency, and the opportunity for every organization”, 5 May 2026. Survey by Edelman Data x Intelligence of 20,000 AI-using knowledge workers in 10 countries, Feb to Apr 2026; and a Microsoft-led study of 1,800 workers (Jul 2025).	https://www.microsoft.com/en-us/worklab/work-trend-index/agents-human-agency-and-the-opportunity-for-every-organization Read 16 Oct 2026
Fortune, on a McKinsey survey. “The companies getting the most from AI are rethinking how work gets done”, 2 Sep 2026: a McKinsey survey of 1,719 professionals and business leaders. McKinsey's own report could not be opened on 16 Oct 2026, so the plan quotes Fortune's report of it.	https://fortune.com/2026/09/02/companies-getting-the-most-from-ai-rethinking-how-work-gets-done-cfo/ Read 16 Oct 2026
Gartner. Press release, 25 Jun 2025: “Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027”.	https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027 Read 16 Oct 2026
Anthropic. Help centre: Change the model, effort, and thinking settings.	https://support.claude.com/en/articles/8664678-change-the-model-effort-and-thinking-settings Read 16 Oct 2026
UK National Cyber Security Centre. Blog: Managing the cyber risk of agentic AI, 20 Aug 2026.	https://www.ncsc.gov.uk/blogs/managing-the-cyber-risk-of-agentic-ai Read 16 Oct 2026
European Union. Regulation (EU) 2024/1689, the AI Act: Article 5(1)(f), Article 50(4), Article 113 and Annex III point 4.	https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32024R1689 Read 16 Oct 2026
European Union. Regulation (EU) 2026/1744, the Digital Omnibus on AI, 8 Jul 2026, Official Journal 24 Jul 2026: Article 1 points (5), (39) and (40), and Article 4.	https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32026R1744 Read 16 Oct 2026
Illinois General Assembly. Public Act 103-0804 (HB3773), amending the Illinois Human Rights Act, Section 2-102(L); effective 1 Jan 2026.	https://www.ilga.gov/Legislation/publicacts/view/103-0804 Read 16 Oct 2026
UK National Cyber Security Centre. Managing the cyber risk of agentic AI, 20 Aug 2026: the line on the leash was checked against it.	https://www.ncsc.gov.uk/blogs/managing-the-cyber-risk-of-agentic-ai Read 16 Oct 2026

Survey figures name who ran the survey, how many people took part, and when it was published or reported. Microsoft and McKinsey sell AI services, and their surveys are self-reported.

What next

FREE, THE WHOLE SET

The Team Pack

Every piece this plan names, free and in one place: the Team AI Policy Kit, both Workshops in a Box, the AI Fact-Check Kit, the Agent Playbook and this plan, run as one 90-day programme.

FREE, IN THE FIELD GUIDE

Level 6: Advanced and by role

Free lessons for managers on leading the change, team rules and measuring value, each checked against its sources, published as they are ready, from neosaipien.com

If you would rather not do it alone

I am Samson Mathew, and Neosaipien is my AI training and consulting practice. Where this plan says “I”, it means me.

LEADERSHIP BRIEFING

For the leadership team: 90 minutes, the alignment session run with you, so the room leaves with its ten answers.

EXECUTIVE COACHING

For the leader: six one-to-one sessions on your own use of AI, and the decisions on page 4.

TEAM WORKSHOP

For each pilot team: half a day, hands-on, run live on the team's own real work.

WORKFLOW AUTOMATION

For the jobs the picker marks “automation”: I can build the fixed steps with your team.

AI IMPLEMENTATION

I can run the 90 days with you: the owner's work, the reviews and the day-90 report.

Book a call

Twenty minutes on where your organisation is, and which of these would help: neosaipien.com

neosaipien

About. Neosaipien is an AI training and consulting practice: the Leadership briefing, Executive coaching, the Team workshop, Workflow automation and AI implementation. Every kit is checked against its sources before it is published. neosaipien.com / samson@neosaipien.com

October 2026 edition. Checked on the vendors' and lawmakers' own pages on 16 Oct 2026. These products change monthly. Check the settings in your own tool. Not legal advice. Neosaipien is not affiliated with Anthropic, OpenAI or Google.

Free from Neosaipien under the free licence, Version 2, in the LICENCE.txt file that comes with it. Use it as much as you like and change it to fit your team. Hand your changed version to people inside your own organisation, team, class or group, keeping the credit and the date and saying it was changed. Share unchanged copies with anyone, free of charge. Do not sell it, and do not publish a changed version outside your own group.