



A PRACTICAL GUIDE FOR LEADERS

The agentic leader playbook

How to build and work with AI agents yourself, then take your organisation with you. For leaders of teams, functions and companies.

Why leaders go first

Redesign the work

A first month for your team

Governance at both ends

Bots as colleagues

Measuring what matters

The six-part agent prompt

Quality checks (the harness)

Ben Legg

Founder, Digital People
Former COO, Google Europe
benlegg@digitalpeople.online

October 2026

Read it in 30 minutes. Lead from it for a year.

AI in 2026: from content to execution

For two years most of us used AI to make content: summaries, drafts, research. From 2026 onwards the value shifts to execution, where AI agents carry whole tasks end to end. This guide is for the people who have to lead that shift, not just use it.

2024 TO 2025

From: content

- Drafting emails
- Summarising information
- Sounding board
- Market research
- Translating between audiences



2026 ONWARDS

To: execution

- Meeting prep and follow-up
- Recurring tasks and briefings
- Data management and report generation
- Project coordination
- Inbox triage, drafting and sending

An agentic business is one where agents do the work and people set direction, brief and check. Teams form around tasks rather than fixed reporting lines, and agents hand work to other agents. It starts with one leader building one agent.

Working on your own productivity? Everything you need is on screen at digitalpeople.online/agent-builder: a step-by-step Claude setup guide and copy-and-paste agent guides. This PDF adds what a leader needs on top: the case, the organisation, the governance. Page 20 is a glossary.

IN THIS GUIDE

01	Why agents	3
	The untrained intern, what an agent actually is, six reasons agents win	
02	Lead from the front	4
	Why leaders go first, redesigning the work, taking your team with you	
03	Governance and the organisation	7
	Risk at both ends, owning what your bots say, bots as colleagues, making your company agentic, measuring what matters	
04	Build your own	10
	Agents worth building, the six-part prompt, where an agent lives, Project or Skill, connectors, voice, model choice, when it breaks	
05	Quality	18
	The AI harness: how to know an agent is good enough to trust	
06	Your commitment	19
	For you, by Friday. For your team, this quarter	
07	Glossary and the questions people always ask	20
	Terms in plain English (20), then the 10 questions every newcomer asks (22)	

Why agents

Generic AI is useful. An agent you have trained is a compounding teammate. The difference is the whole point.

The problem: generic AI is an endless line of untrained interns

GENERIC CHAT

New untrained intern. Every morning.

Twenty minutes explaining who you are, who your audience is, the voice you want, the things they should never do. Then the actual work. Tomorrow morning, same conversation.

YOUR AGENT

Train once. Remember forever.

Same job, same rules, same memory, the morning after, and the morning after that. The intern you only have to train once. A compounding teammate.

What is an agent?

An AI you have trained for one specific job. In practice it is a page of written instructions, plus the files and memory Claude keeps alongside them. Claude reads that page every time, so you explain the job once, and it does that one job a little better every time you use it.

That is all it is. No code, no engineers. If you can describe the job to a competent new hire, you can write the page. Page 11 gives you the template. The cost is an hour of setup and a 20-minute Friday habit (page 17).

Six reasons agents win

Each element is useful on its own. Together they compound.

Specialisation

One job done well. Never distracted.

Memory

It remembers your style, your audience, your people.

Voice

Learns to sound like you, not like an AI.

Boundaries

Rules set once. They stick.

Reach

Wired into your calendar, email and files (page 15).

Compounding

Tightened weekly. Better every year.

Why leaders have to go first

You will be asked to approve the AI strategy, the budget and the risk. You cannot judge any of them well from a slide deck.

5 hours a week

is Ben's minimum for a leader who wants to be ahead, not average: building, managing and improving agents. The average executive manages about 2. Only 7% did more than 5.

Yotzov, Bloom et al., NBER w34836, revised Oct 2026; 7% from its early 2026 survey.

80% and 37%

80% say AI has improved their own productivity. Only 37% say it has contributed to EBIT.

McKinsey, The state of AI in 2026, August 2026.

About 6%

are AI high performers, with 5% or more of EBIT from AI.

Same McKinsey survey.

We read the gap between the 80% and the 37% as a leadership gap. Individuals get faster on their own; the organisation only gets better when someone redesigns the work. That someone is you.

The agentic leader: 5 shifts in your job

- 1 **Be an AI power user.** At least 5 hours a week, building and working with agents on your own work. You cannot design work you don't understand.
- 2 **Redesign the work.** Which tasks are human, which are agent, and how do the two work together? (page 5)
- 3 **Make the plan visible.** A change programme your team can see, with the destination clear.
- 4 **Take the team with you.** Teach it, learn together, and correct course as you go. (page 6)
- 5 **Make it systematic.** A route for good ideas to become firm-wide improvements, and the habit of writing every lesson back into the agent.

Shift 1 is the one you can't delegate.

What one hour of building teaches you

- What an agent gets wrong, and why.
- What data it really needs.
- Where the risk actually sits.
- Better questions for your CIO and board.

Your people copy what you do

Use an agent openly, show your drafts and admit what went wrong, and you give everyone permission to try. Hand AI to IT, and they learn it is something done to them.

Two org charts

Org design is one of the biggest drivers of productivity, and AI changes what it means. You inherited one org chart. You now have to design a second one alongside it.

THE ORG CHART YOU INHERITED

People, arranged by grade

Hired, promoted, appraised and reassured. You know how to run this one. It has not stopped being your job.

THE ONE YOU NOW DESIGN

Humans and agents, arranged by outcome

Each outcome (win the work, do the work, keep the customer) has a human owner, the agents that do the steps and a human reviewer. It is never finished: it has to be designed and continually upgraded.

Sort every task into three piles

Do this with your team, one process at a time. It is the fastest way to find where agents belong.

Green

AI does it, I check

High volume, low stakes, easy to reverse: sorting, summarising, first research, routine updates.

Amber

AI drafts, I decide

Anything that goes out under your name or changes money, plans or priorities.

Red

I do it

Judgements about people, relationships that matter, and anything you could not take back.

The compounding-error test

If each of 3 steps in a chain is 70% reliable, the chain is only 34% reliable ($0.7 \times 0.7 \times 0.7$). Before you link agents together, make each step reliable, or put a human check where the errors would compound.

Build, Buy, Borrow, Bot: one decision per task

For every recurring task, decide who or what does it. **Build** it in your own team.

Buy it from a vendor. **Borrow** fractional or contract talent. Or hand it to a **Bot**, an agent you have trained. The mix of full-time, fractional and agent is now a design choice, not a default.

Orchestrating, not doing

People set direction, brief and check, while agents do the doing and hand work to other agents. The question moves from how many people you need to which skills. The leader's job shifts from managing tasks to designing the system that does them.

A first month for a leadership team

Gartner expects over 40% of agentic AI projects to be cancelled by the end of 2027, mostly on cost, unclear value or weak risk controls (press release, 25 June 2025). A short, visible plan that starts with the leaders, with one number per agent, is the antidote.

1

Week 1

You build 2

Set up Claude with the online setup guide, then the calendar brief and the email assistant. About 3 hours, on an account IT has approved.

2

Weeks 2 and 3

Team builds 1 each

Each member of your leadership team picks a guide, builds it on their own work and shows the team what changed.

3

Week 4

Keep the best 3

Turn them into shared agents, each with a named owner, a short job description and data rules.

4

After that

Roll out with rules

Approved tools on a company account, short rules people remember, and a monthly review of what is running.

Make it stick

- ☐ **Baseline first.** Write down how long the task takes today, before anyone builds anything.
- ☐ **One number per agent.** Minutes saved, turnaround time or errors caught. One, not five.
- ☐ **Check in at 10 to 14 days.** Most agents are good after 3 to 5 runs with clear feedback. Most drop-off happens before that.
- ☐ **Ask what you will stop doing.** Time saved only counts once you decide where it goes.
- ☐ **Run “show me yours”.** 30 minutes, everyone demos one agent.
- ☐ **Say what happens to jobs.** Tell people early where saved time goes. Silence is read as cuts, and in Europe consultation duties may apply.

Design for mixed ability

- Some of your team built an agent last week. Others have never opened Claude. Both are normal.
- Make the first 5 minutes hands-on: everyone does one real task from their own inbox with AI, before any teaching.
- Pair your fastest builder with your most sceptical senior person.
- Setup is the most important part. Check access before the session, not during it.
- Judge the decision the agent helped make, not the cleverness of the agent.

The delegation ladder

Most leaders are on rung 1. The guides take you to rung 4 in a week. Rung 5 is where the organisation changes.

1

Ask

You ask AI a question, and start from scratch every time.

2

Brief

A standing brief in a project, so it remembers who you are.

3

Connect

It reads your diary, inbox and files, and drafts from them.

4

Schedule

It runs before you ask, and you review the results.

5

Supervise

You manage a team of agents, each with one job and an owner.

Make every task a learning opportunity, for the agent and for you.

How to think about AI risk

You can get governance wrong at both ends. In Ben's experience advising 50-plus companies, firms that ban AI rarely have fewer incidents; they just see fewer, because caution breeds shadow AI: staff using AI on personal accounts where nobody can see it.

Too tight

Ban AI, or say nothing for ages

- Your best people use AI anyway, on personal accounts. Shadow AI.
- You have not removed the risk, only your visibility of it.
- Your most capable people learn fastest, and they learn outside your walls.

The middle

Where the incidents are fewest

- **Tools.** Provide best-in-class AI tools.
- **Training.** For productivity and for governance.
- **Simple rules.** Short enough that people remember them.
- **Build freely, for yourself.** Personal agents may read, but should not send, pay or publish without a human.
- **Approval before scale.** Formally review every agent before it goes firmwide.

Too loose

Any tool, any data, no guidance or approvals

- Hundreds of badly built agents, using data in random ways.
- Nobody can tell you what is running, or where the risks are.
- The first incident becomes the reason to swing to "too tight".

You own what your bot says

The chatbot that wrote its own policy

Air Canada's website chatbot told a grieving customer he could claim a bereavement fare after travel. The real policy said before. In February 2024 a Canadian tribunal held the airline to the chatbot's answer. Air Canada argued it was not liable for what its chatbot said. The tribunal said this in effect treated the chatbot as a separate legal entity, and called it "a remarkable submission".

Moffatt v. Air Canada, 2024 BCCRT 149. Award: C\$812.02.

The lessons

- Nobody wrote that policy. It still bound the company.
- The money was never the risk. The risk was arguing that the machine was responsible.
- Regulators expect human review to be real. The UK ICO says it must be meaningful, not a token gesture.
- **The rule:** keep a human on every decision about a person. Free up the routine.

Where the data actually goes. Your data is stored in the US on every tier, and by default a prompt may be processed wherever is fastest, so EU personal data leaves the EEA either way. There is no European residency for the Claude app: Enterprise can pin processing to the US and adds audit logs, which is a stronger contract, not a European one. Keep special-category data, meaning health, ethnicity and the like, out of prompts until legal have signed it off. *Checked September 2026.*

Your newest colleagues are bots

Agents are joining faster than your policies were written for them. Every leadership team now needs answers to 8 new questions.

1. Who is the manager?

Every agent needs a named human owner. If nobody can say whose it is, it should not be running.

2. When is “my bot” “our bot”?

At the second user. That is the moment it needs approval, a short job description and data rules.

3. Does it get a performance review?

Yes, and often. Its users review it, and an agent that underperforms gets fixed or retired.

4. How do you reward the builder?

They have created a new colleague. Recognise it, or your best builders will stop sharing.

5. The builder leaves. Does the bot leave?

A ghost agent with no owner keeps running. Make agent handover part of every leaver checklist.

6. Who owns the agent?

The prompt, its memory and the data it built up. Your employment and supplier contracts probably do not say.

7. Is it on the org chart?

A small team of people can now run many agents. Show them, so you can see what the team really does.

8. Does it need a work permit?

Built in London, German data, US servers. Where does it work? Ask legal and tax before you scale it.

Making agents official company bots

Anyone can build for themselves. Before an agent becomes an official, shared company bot, it should pass 4 tests.

Importance of the problem being solved

Is this a real bottleneck for real people, or a clever demo?

Uniqueness of the solution

Does it do something the existing tools and vendors do not, or do badly?

Financial impact of the agent

Hours saved is a vanity metric. Decide where the hour goes before you save it.

Appropriate compliance safeguards

What data does it read, what can it change, and who checks it? Page 18 is the checklist.

8 steps, then measure what matters

Personal agents get you fluent. Company-wide execution is a different job, and it is yours. 8 steps, in 2 halves.

Company-wide execution: how to get started with AI

1

Make all leaders fluent daily users

And be seen using them.

2

Define your full-potential AI vision

What the business looks like when agents do the work.

3

Gather, clean and protect your data

Agents are only as good as what they can read.

4

Ensure appropriate governance

See page 7. Wrong at both ends is possible.

How to win: make your company agentic

5

Prioritise the top 3 to 5 AI initiatives

And resource them appropriately.

6

Adopt Build, Buy, Borrow, Bot for each task

Decide who or what does every task.

7

Build and deploy specialist agents

One job each, done well.

8

Integrate your multi-agent system

Agents cooperating with agents.

Measure what matters

6 numbers for the board

- **Adoption by layer:** executives, managers, front line. Leaders first.
- **One number per shared agent,** against its baseline.
- **Quality:** errors caught, and errors that got through.
- **Where the time went:** customers, growth, or simply gone?
- **What is running:** every shared agent, with its owner.
- **Cost:** licences and usage per shared agent, against the value of the time it frees.

2 early examples worth knowing

Morgan Stanley gave its wealth advisers an AI assistant over its own research. Over 98% of adviser teams use it (OpenAI case study, 2024).

IKEA moved routine customer questions to its chatbot, Billie, then retrained 8,500 call-centre staff as remote interior design advisers (Reuters, June 2023). Growth, not just cost.

Copilot, Claude, or both?

The real differences between platforms are less about the model and more about where your data lives, who controls the switches and how the meter runs. If your company runs on Microsoft 365, start where your data and your admin controls already are. Learn what a good agent looks like here, then ask IT which parts your platform can do today.

Popular agents to consider

Start with a task, not a strategy. Some agents take on recurring work you would happily hand over; others exist for a flash of input or inspiration. 6 for work and 6 for life, all real, all built in under an hour. Pick the one or two that would change your week.

Six common agents for your work Each card reads: what you give it, then what you get back

Email inbox triage

Every unread email sorted four ways: urgent, reply, read later, unsubscribe. Draft replies created for you to review.

Meeting prep

A daily one-page brief per external meeting. End-of-day follow-up drafted.

Group update

Voice note in. 250-word update email out.

Pipeline management

Daily updates to the sales pipeline: contract status, meetings, emails.

Project management

Research, scoping, deliverables, KPIs, meetings and updates, kept current.

KPI synthesiser

Folder in. Three-page synthesis out, cited.

Six agent ideas for your life

Sleepyhead

Story seeds in. Calming 700-word bedtime story out. Built in 20 minutes for Ben's 2 young children.

Holiday architect

7-day family itinerary: plan, budget, bookings, checklist.

Gift curator

Name, age, budget in. Three gifts out.

Health and fitness coach

Last session in. Tomorrow's session out.

Recipe magician

Photo of the fridge in. Three dinner ideas out.

Wedding planner

Compare vendors, track the budget, chase the deadlines.



Sleepyhead in action. Story ideas go in by voice; a 700-word bedtime story comes out, with two or three decisions for the children to make along the way.



The recurring cast. 'Super Jasper' and 'Pixie Pippa' are written into the agent's instructions, so every story knows who the heroes are. Page 11 shows the prompt.

Why the life ones matter

A bedtime story agent does more to make agents feel achievable than any B2B example. Build one for your family first if you like. The muscle is identical, and the stakes are low enough that you will finish it.

The agent prompt

6 parts. Use them for every agent you build. This is the single most valuable page in the guide, because it is copyable.

The six parts	Example: storyteller agent	Example: email triage manager
1 Role Who is the agent?	“ You are Pippa and Jasper’s storyteller ”	“ You are my email triage manager ”
2 Goal The one job, in a sentence.	“ Create entertaining bedtime stories and discussions ”	“ Help me respond efficiently and effectively to all emails ”
3 Inputs What you give it.	“ ‘Super Jasper’, ‘Pixie Pippa’, our life ”	“ You are connected to my email (page 15 explains connections) ”
4 Constraints Rules. Always. Never.	“ UK English; educational and funny; make them think ”	“ UK English. Warm but brief. Never long dashes. Never invent. ”
5 Outputs Exact format you want back.	“ Stories for 5 to 10 minutes, with 2 to 3 key decisions per story ”	“ Flag spam to unsubscribe; draft every reply for my review ”
6 Failure modes When to stop and ask.	“ If in doubt, we can discuss ideas ”	“ If in doubt, flag with questions ”

How to write yours

- **Talk it, don't type it.** Describe the job out loud as if briefing a new hire (page 15 shows you how). 6 parts, 6 short lines. That is the first draft.
- **Constraints are where the value is.** Your never-do list is what stops it sounding like an AI.
- **Failure modes make it safe.** An agent that stops and asks beats an agent that invents.
- **Tighten it weekly.** Every time it gets something wrong, add one line. That is how it compounds.

Space for your own six lines

On paper, draft here. On screen, open a note and do the same. Either way it then goes into a project (the online setup guide shows you where).

An agent is a file. It has to live somewhere.

Four homes. Pick each agent's home based on what level of support you need, how often you run it and who else needs it. Strictly, Home 2 is where the rules that follow every agent live, rather than an agent itself.

HOME 1

A chat project

One job, one workspace, its own files and its own memory. **Thinks with you, but doesn't execute.**

HOME 2

A Claude Skill

A rule that follows you everywhere. Your voice, your formatting, your never-do list.

HOME 3

A Cowork project

When it needs to **do** the work, not draft it. Runs for hours; online-only tasks carry on when you close the laptop.

↓ sits within

The rule of thumb

- **Thinking or a draft: a chat project. An action: Cowork. If in doubt, Cowork.**
- **Promote it to a Skill** when you find yourself repeating the rule across projects.
- **Schedule it when it earns the right.** A scheduled task sits inside Cowork and inherits the project's instructions, files and memory.

HOME 4

A scheduled task

Lives inside a Cowork project. When it should run without you asking. Daily, weekly, or every weekday at seven.

Which home for which agent?	
Meeting prep brief, every morning	Scheduled task, inside a Cowork project
Your writing voice	Skill
Board pack thinking partner	Chat project
Inbox triage that drafts replies	Cowork project
Bedtime storyteller	Chat project (or the equivalent in Google Gemini)

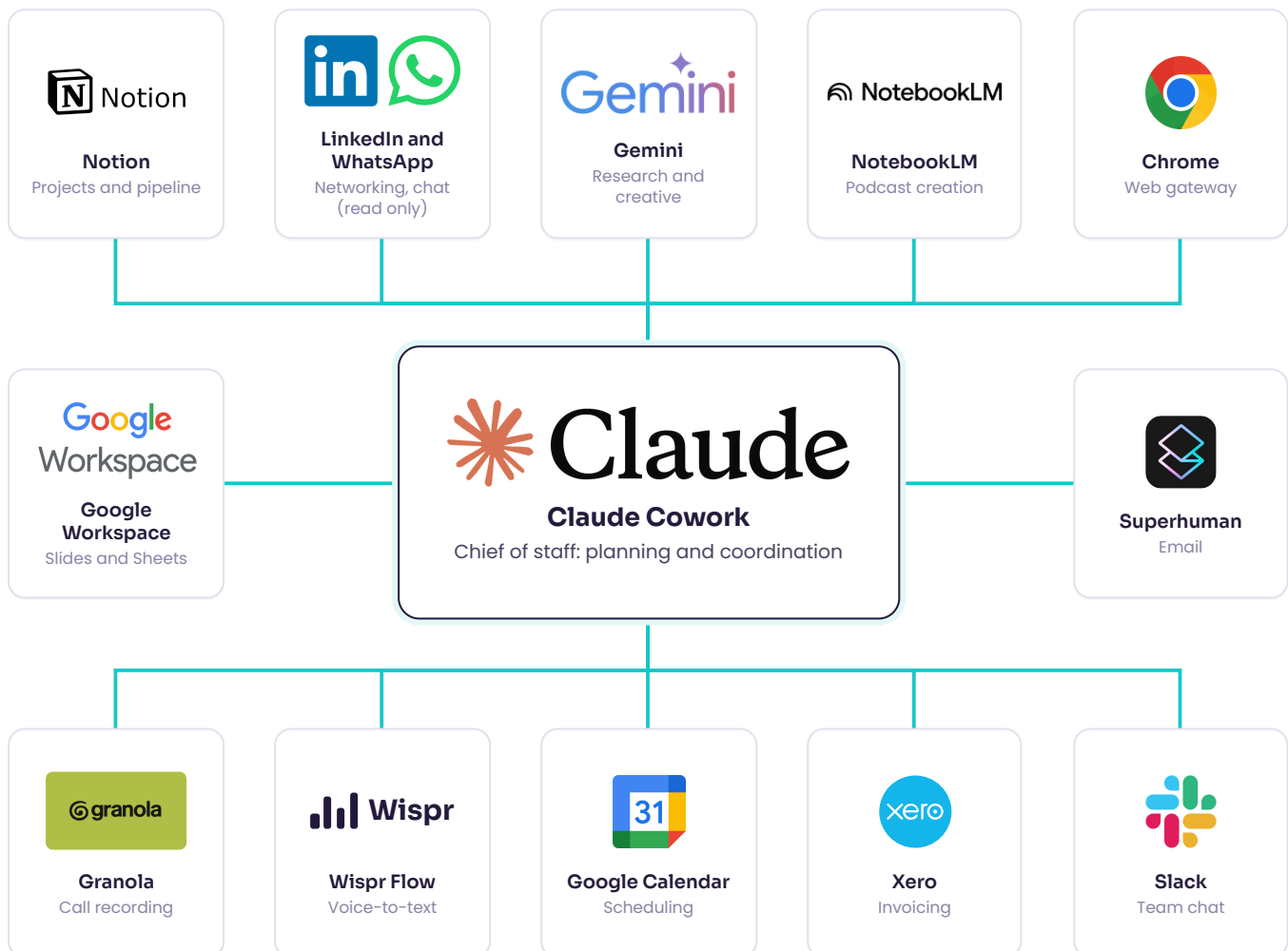
Ben's own stack is on the next page

12 tools wired around Claude Cowork, each doing one job. In daily use, and yours to copy.

Ben's Claude-centric agentic stack

Snapshot: September 2026

This is not a theory. Claude Cowork sits in the centre as chief of staff, planning and coordinating, and is wired into 12 tools where the work actually lives. Each tool does one job. It is in daily use, and you are welcome to copy it.



Where to start

Email, calendar and voice first (page 15): they teach Claude who you are and double your speed in a week. Then your document store. Then the rest, one tool at a time.

Dated for a reason

Ben changes something in this stack roughly once a month, so check the date above. The pattern is what lasts: one orchestrator, many specialists, each connected once.

Project or Skill?

Both hold instructions. One stays in its room, the other follows you everywhere. Getting this right is the difference between an agent that works and a pile of pasted prompts.

PROJECT

Local. Stays in its room.

Use it for the work itself.

Instructions, files and memory that apply to **one job only**. Your board pack project does not know about your recruitment project, and that is the point.

Manage the job here: the brief, the source files, the outputs, the memory of what worked last time.

SKILL

Global. Follows you everywhere.

Use it for how you want everything done.

One rule, available in **every project and every chat**. Ben's 'write like Ben' skill was written from a year of his sent email; 20 to 30 emails is enough to start. It runs whether he asks for it or not.

Voice, formatting, house style, your never-do list, how you like a brief laid out.

And underneath both: your Instructions. *Settings, then Instructions for Claude* holds the rules that apply to everything you ask: tone, language, formatting, your never-do list. "Always write in UK English" is an **Instruction**. "How I write a client proposal, step by step" is a **Skill**. The proposal you send on Friday is a **Project**. Instructions apply inside projects too, so there is no need to paste them in; a project's own instructions sit on top, for that job.

Rule of thumb: if you have typed the same instruction into three different projects, it should have been a Skill.

Three skills almost everyone should build first

- 1 **Your voice.** Paste in 20 to 30 of your own sent emails (strip client names first if you are on a personal plan) and ask Claude to describe how you write: greetings, sign-offs, length, what you never say.
- 2 **Your formatting.** How you like documents, briefs and summaries laid out. Date formats, headings, length.
- 3 **Your never-do list.** The short list of things that make you wince. It is the cheapest quality control you will ever build.

The worry: "people will know it was AI"

This is the most common fear people voice, and the answer is a Skill written from your own emails. Once it knows how you greet people, how you sign off and what you would never say, the output is yours in every way that matters. Review before sending, always. But the drafting is no longer the bottleneck.

Where to find it: Customize, then Skills, then the plus sign, then Create skill. On the desktop app you can also record a skill from the composer.

MCP is like a plug socket

MCP (Model Context Protocol) is the industry-standard plug that lets Claude connect to other apps. Claude calls these connections "connectors". You will see the letters in menus and never need to understand them. Connect in this order.

CONNECT FIRST

Calendar, email, Chrome

These teach Claude who you are and add value from day one. Chrome is the exception to "no mouse": Claude in Chrome does drive your browser, in a tab you can watch. Read its safety guide first.

CONNECT NEXT

Where your work lives

Google Drive or SharePoint, Notion, Slack, Granola, HubSpot and so on. Check whether each connection can only read, or can also write and send. Read-only first.

NOT YET

Your bank. Your WhatsApp.

Not because it will not work. Because the risk is higher. Claude only sees what you can see in a connected app, and only when a task asks. Earn the confidence first.

On a company account there are two approvals, not one

1. Your admin turns the connector on.

Organisation settings, then Connectors, then the connector, then Managed authorisation.

2. Then you authorise it yourself.

Settings, then Connectors, then Connect. Miss the first and the second silently does nothing, and it looks like the product is broken.

You type at 40 words a minute. You talk at 130.

A Stanford study (Ruan et al., 2016) found dictation on a phone three times faster than typing on it, with fewer errors. On a keyboard the gap is smaller but still large. Use your voice for almost everything. The awkwardness lasts about two days. Then you will not go back.

FOR MOST APPS AND WEBSITES

Wispr Flow (a dictation app)

Free tier is 2,000 words a week, which is plenty to be convinced. Push to talk, works in every app on Mac and Windows.

WHILE ON CLAUDE ITSELF

The microphone in the app

Already there. Nothing to set up. The best way to brief an agent while walking to a meeting.

IF YOU CAN'T DOWNLOAD WISPR

Press Windows + H

Already on your machine. Real punctuation. No admin rights, no install, no ticket. On a Mac, press the microphone or fn key twice.

Pick the right model. Then make your agents pick too.

Claude comes in four versions, called models. Think of them as grades of staff at four price points. For a big task Claude often splits the work among helper copies of itself (sub-agents), and each helper can run on a different model. So your own default matters less than what your agents choose for their helpers.

Fable	Big, open-ended jobs where it works things out for itself. Slowest and dearest.	10x
	A real job: work out which three markets we should enter, and why.	
Opus	Judgement, and anything going out in your name	5x
	A real job: draft the client proposal you are about to send, or the board paper.	
Sonnet	Everyday work. Speed and sense, in balance	2x
	A real job: turn these call notes into a summary and three actions.	
Haiku	Fetching, scanning, listing. Volume, at speed	1x
	A real job: pull the last 20 emails from this client and list the dates and subjects.	

Relative price per token at published rates, September 2026. The ratio moves as versions change; re-check it before you rely on it.

The instruction to paste into your settings

Settings, then Instructions for Claude (page 14). It applies to every chat and Cowork task you start. In plain words: cheap models for gathering, the good model for judging and anything that leaves the building, and never let a cheaper model mark a dearer one's work.

When you spawn subagents, match the model to the job: gathering, scanning, extracting and listing run on Haiku or Sonnet. Judging, synthesising and drafting anything I will send run on the session model or better. Never grade work with a cheaper model than produced it.

The £200 lesson

£200 in 2 days, on pay-as-you-go pricing, was what it took Ben to notice he had assumed cleverer was always better. It was not. It was just slower and dearer. **Check your setting in week one**, not in month three when the bill arrives.

50 to 75% lower cost

For the identical job, against every sub-agent left on the default model. Illustration: 11 workers, 10 gathering and 1 judging. All on Opus costs 55 units of cost. Move the gatherers to Haiku and it costs 15.

Why the rule has two halves

Gathering is checkable: you can see whether the ten things came back. Judgement is not, which is why judgement never moves down the ladder. And never let the cheaper model be the one that marks the work.

On a monthly subscription you have already paid, so the same mistake shows up differently: paid plans have a usage allowance that resets every five hours and every week, and the dear models eat it fastest. Everybody hits it eventually, usually at four o'clock on a Thursday. This is how you stop. Bonus: files kept in a project are cached and cost almost nothing on reuse.

When it breaks, it is almost never the AI

Almost every failure in the first month is setup, permissions or context. Find your symptom on the left.

What you see	What it usually is
You cannot get into your work email to connect it	A lost password, and a reset code going to an address you closed. Fix it before you start, not mid-task.
The connector says it is on, but nothing happens	Your admin enabled it, you never authorised it. Two approvals, not one. Page 15.
It gives a confident answer that is wrong	You gave it no instructions to check. Tell it what good looks like and how to check quality. Page 18.
It forgot everything you told it last week	You were in a chat, not a project. Ordinary chats keep only a light general memory; a project keeps its instructions and files every time.
The scheduled task produces poor work	You created it outside the project, so it starts each run without the project's instructions, files and memory. Recreate it from inside.
It keeps stopping to ask permission	Each Cowork task has an approval setting: Manually approve, Automatically approve, or Skip all approvals. Use Automatically approve for reading and drafting; never Skip for anything that sends or pays. You are not being careful, you are being interrupted.
It sounds like an AI, not like you	No voice skill. Build one from your own sent writing. Page 14.
It is slow and eating your usage limits	Everything is running on the most expensive model, including the sub-agents doing the fetching. Page 16.

The Friday habit: twenty minutes

Every Friday afternoon, open the agent you use most and ask it three questions: What did you get wrong this week? What did I have to correct twice? What one line should I add to your instructions? Add the line. That is the entire maintenance routine.

Learning AI doesn't have to be over-engineered

The Legg family runs 60-minute AI brainstorming on Zoom, two or three a month: joint problem solving and sharing what each person has learnt. Do the same with your team. Ten minutes of "show me yours" beats an hour of slides.

Before you blame the model, ask: right account, right home, inside the project, both approvals done, instructions to check? 9 times out of 10 the answer is in that list.

How do you know it is any good?

Everyone builds the agent and forgets the harness: the set of checks around it that catch mistakes before they matter. Four questions. Each answer is something you already know how to build.

QUESTION 1

Does it know what good looks like?



IF NO

Write the standard into it

Nothing else on this page works without it. Decide what a good answer looks like before you build.

QUESTION 2

Do you want that standard on every job?



IF YES

Put it in a Skill

It then follows you into every project and every chat, whether you remember to ask or not.

QUESTION 3

Is it a judgement call, not a right answer?



IF YES

A second agent marks it

Fresh eyes, a different prompt, a standard written first. In practice: a second project whose instructions begin "You are the reviewer"; paste the output in. Never the one that wrote it.

QUESTION 4

Could you take it back if it were wrong?



IF NO

It stops. You look.

Sending, paying, publishing, deleting. When all four have an answer you are happy with, let it run.

Why this order?

An agent marking its own homework gives itself an A. Peer-reviewed research (Panickssery et al., 2024) shows AI graders favour their own output. The first three questions are you asking. The last one is the only guarantee.

Read is not write. Draft is not send. Letting Claude read your email is not the same as letting it send. Most permission worries dissolve once you separate the two.

Lines to remember

- One check is theatre. Three is a harness.
- If your agent has never stopped and asked you something, it is not being careful. It is being confident.
- Put the human gate where the money is, not where the anxiety is.
- Chain three agents at 70% each and you get 34% end to end. Reliability multiplies down a chain. (Ben's arithmetic, not a study.)

What will you do differently?

Not a strategy. A task, with a name, by Friday. Then a short list for your team this quarter.

For you, this week

- ☐ Work through stage 1 of the online setup guide: the right plan, your calendar and email connected
- ☐ Build the calendar brief, then the email assistant, from the guides at digitalpeople.online/agent-builder
- ☐ Use your voice instead of typing for one full day (page 15)
- ☐ Write the six-part prompt for one agent of your own (page 11)
- ☐ Block 5 hours a week in your diary for building and improving agents, starting with a 20-minute Friday review (page 17)

For your team, this quarter

- ☐ Run the first month with your leadership team (page 6)
- ☐ Sort one core process into green, amber and red (page 5)
- ☐ Agree your simple rules for personal agents versus official company bots (pages 7 and 8)
- ☐ Give every shared agent a named owner and one number (page 9)
- ☐ Name your company's top 3 to 5 AI initiatives (page 9)
- ☐ Add the four harness checks to any agent that sends, pays or publishes (page 18)

The one agent I will build first

Its name, what it does, and the one number that will tell you it works.

The one thing my team will have done in 30 days

DISCUSSION PROMPTS FOR YOUR TEAM

- Which of us will build first, and show the others?
- Which process do we sort into green, amber and red first?
- What are our rules, short enough to remember?
- Who owns each shared agent, and how do we review it?
- Where will the time we save go?

KEEP IN TOUCH

Ben Legg

Founder, Digital People. Former COO of Google Europe, ex-McKinsey and 3 x PE-backed tech CEO. Advises CEOs, boards and investors on building AI-first organisations.

benlegg@digitalpeople.online

Happy to hear which agent you built. Send me a line.

Glossary

25 words and phrases you will hear in the boardroom and see in the menus, in plain English. 2 pages.

AI agent

A version of Claude given one job, a page of written instructions and access to the tools it needs (calendar, email, files). Unlike a chat, it carries a task through to a finished result and checks in with you at agreed points.

Agentic

Work, or a business, where AI agents carry out tasks and people set direction, brief and check. "Going agentic" means moving from asking AI questions to handing it jobs.

Model (LLM)

The underlying AI engine that reads and writes. Claude is a family of models; ChatGPT and Gemini are rival families. LLM stands for large language model. You will never need the phrase.

Haiku, Sonnet, Opus, Fable

The four Claude models, from cheapest and fastest to dearest and most capable: roughly 1x, 2x, 5x and 10x the cost. Haiku for fetching and listing, Sonnet for everyday work, Opus for judgement, Fable for long open-ended jobs.

Prompt

What you tell Claude to do, spoken or typed. The six-part agent prompt on page 11 is a prompt you write once and keep.

Instructions

The standing rules under everything you ask Claude to do: tone, language, formatting, your never-do list. Set them in Settings, then Instructions for Claude. They apply to every conversation, projects included; a project's own instructions sit on top of them for that one job. Page 14.

Project

A workspace inside Claude for one job, holding its instructions, its files and its own memory. Nothing in one project leaks into another. Chat projects are for thinking and drafting; Cowork projects are for doing.

Skill

A reusable rule Claude applies in every chat and project whether you ask or not: your writing voice, your formatting, your never-do list. Technically a short text file. On the desktop app you can also create one by recording yourself doing a task.

Connector (MCP)

The plug that links Claude to another app (Outlook, Google Calendar, SharePoint, Slack) so Claude can read from it and, if you allow, act in it. MCP, Model Context Protocol, is the industry standard behind the plug. Ignore the letters.

Cowork

Claude's mode for delegated work rather than conversation. You brief a task; Claude plans it, uses your files and connectors, shows you the plan and asks before doing anything that matters. Fullest on the desktop app; also on web and phone. Paid plans only.

Scheduled task

An agent set to run on its own at a fixed time (every weekday at seven, every Monday). Lives in Cowork and runs on Anthropic's computers even when yours is off, unless it needs files that live on your machine.

Sub-agent

A helper copy of Claude that the main agent starts to take on part of a big task, often several at once. Each can run on a different model, which is why page 16 matters.

Token

The unit AI is metered in: roughly three-quarters of a word. Everything Claude reads and writes is counted in tokens, and your plan's usage limit is a token budget in disguise.

Continued on the next page.

Glossary (continued)

Usage limit

The allowance on paid plans, measured over five hours and over a week. Hit it and Claude pauses until it resets. Dear models and long chats eat it fastest; files kept in a project are cached and cost little on reuse.

Context

Everything Claude can see while working on your task: the instructions, the files, the conversation so far and what the connectors return. An agent "with no context" is working blind.

Memory

What Claude keeps between sessions. Each project has its own memory; Claude also keeps a light general memory across ordinary chats on personal plans (off by default on Team and Enterprise). You can read, edit, pause or wipe it in Settings.

Harness

The checks you put around an agent so mistakes are caught before they matter: a written standard, a second agent to mark the work, and a human stop before anything irreversible. One check is theatre; three is a harness.

Hallucination

When an AI states something false with complete confidence: an invented figure, quote or meeting. The cure is not a cleverer model but an instruction to check, cite and say "I don't know".

Approval modes

How much a Cowork task may do without asking: Manually approve (asks every time), Automatically approve (works on, asks only for things that matter) and Skip all approvals (never asks). Use the middle one. Never Skip for anything that sends, pays or deletes.

Read versus write

The most useful permission distinction there is. Read: Claude can look at your email. Write: Claude can send it. Give read access freely; give write access only once the harness is in place.

Copilot versus Claude

Microsoft Copilot is Microsoft's assistant inside Office. A company admin can choose Claude as the engine inside some Copilot features. That gives you Claude's brain in Microsoft's product, but none of Claude's workshop: no projects, skills, connectors, Cowork or scheduled tasks.

Plan

The subscription tier that decides what you can build: Free, Pro, Max, Team, Enterprise. Free is for trying; Pro and Max are personal and unshareable; Team and Enterprise let you share projects and skills and give the company controls over data and connectors.

Shadow AI

Staff using AI on personal accounts because the company has banned it or said nothing. The risk has not gone; the company has simply lost sight of it.

Markdown file

A plain text file with light formatting marks (a # for a heading, a dash for a bullet) whose name ends in .md. Skills and many agent instructions are written in it. Notepad or any text editor opens it.

Voice dictation

Speaking instead of typing. Claude has a microphone button; Wispr Flow works in every application; Windows has it built in (Windows key + H) and so does Mac (press the microphone or fn key twice). Roughly three times faster than typing on a phone.

Three phrases you will hear that mean less than they sound

"**Agentic workflow**" is a sequence of agents handing work to each other. "**Multi-agent system**" is the same thing at company scale. "**Orchestration**" is the job of deciding which agent does what, in what order, and who checks. None of them needs a consultant to explain.

10 questions every newcomer asks

The questions people ask Ben, and the ones they search for when nobody is watching. Short answers; the page references carry the detail.

What will this cost me, and can I buy it without IT?

Pro is a personal monthly subscription you can buy yourself in minutes. Team and Enterprise are bought by the company. If your firm has a Claude plan, ask for a seat first; if it does not, Pro is enough for everything in this guide except sharing.

Where does my data go?

To Anthropic's servers, where the model runs. Page 7 covers where it physically sits. On Free, Pro and Max, switch off "Help improve our AI models" in Settings, then Privacy, and your chats are not used for training and are kept for 30 days. Team and Enterprise are never used for training. Connected apps process data on their own terms.

Desktop app, website or phone?

All three run chats and projects. Cowork is fullest on the desktop app, which is also the only place it can read files on your computer and the only place you can record a skill. Start on the desktop.

Chat project or Cowork project? They look the same.

Both hold instructions, files and memory. A chat project thinks with you and hands you a draft. A Cowork project acts: it edits the files, sends the email you approved, fills the spreadsheet. If the output is a thing that happens rather than a thing you read, use Cowork.

Is a Skill an agent?

No. A Skill is a rule that every agent and chat follows. Your voice skill and your formatting skill can run at the same time; if two clash, the more specific one usually wins, and you fix the clash by editing the text.

What happens if a scheduled task fails at 7am?

Cowork shows the run and its result under Scheduled, and the task's output tells you it stopped and why. Nothing is sent or changed silently; check the Scheduled list the first few mornings until you trust it.

How do I stop Claude sending something I have not seen?

Give it read access, not send access, and write "draft, never send" into its constraints. Keep the approval mode on Manually or Automatically approve, never Skip. Page 18 is the full checklist.

Can I move an agent from a chat project to Cowork later?

Yes. The agent is the instructions plus the files. Copy both into a new Cowork project and it carries on; the memory of past chats does not travel, so paste in anything it must remember.

How do I give an agent to a colleague if I am on Pro?

Send them the six-part prompt and the files; they paste both into their own project. What you cannot do on Pro or Max is share a live copy that stays in sync. That needs a Team or Enterprise plan.

Who is responsible if an agent gets it wrong with a client?

You are, exactly as with a junior colleague's draft you signed off. That is why the harness puts a human at the point where money, reputation or data leaves the building, and why "draft is not send" is the rule to remember.

Still stuck? Anthropic's help centre (support.claude.com) answers most set-up questions, and page 17 covers the eight things that usually break.