

FOR OWNERS WHO WANT AI WORKING WHILE THEY SLEEP

# Your First AI Employee

*A non-technical guide to autonomous AI agents for small and medium-sized businesses. What they can do, what they cost, and when to stop.*

---

**AUTHOR**

Ted Milton  
Founder, Expedient AI

**EDITION**

July 2026  
Version 1.0

**FORMAT**

A4 print edition  
12 pages, worksheets included

---

**Expedient AI Ltd**

INDEPENDENT · BRISTOL · UK

## 1 THE DIFFERENCE THAT MATTERS

# An agent is not a chatbot.

*A chatbot answers questions. An agent takes actions.*

Chat with most AI tools and they produce text on screen. Useful, but you still have to do the next step yourself. An autonomous agent is a system that can take actions in the world, with permission. It can send an email, update a record, post a comment, fetch a page, file a ticket. The difference is not the model. The difference is the permissions it has been granted and the tools it can call.

For a small business, the practical difference is between "ask it a question" and "give it a job". Asking questions is what most people have used so far. Giving it a job means describing a task, granting the permissions it needs, and reviewing the work it does.

## Why this matters for a small business

Most of the value of AI in a small business comes from automation, not from questions. The work that takes time is repetitive, structured, and follows a pattern. An agent can do that work between meetings, overnight, while the team is on the train. The owner gets the result, not the typing.

## What an agent is, in one paragraph

An agent is a system that has been given access to other tools, a written description of the job, and a set of permissions. It reads the input, decides what to do, calls the tools it needs, and produces a result. A human reviews the result. The agent is not magical. It is a loop: read, think, call a tool, read the result, repeat, until the job is done or it gives up.

**The trade.** An agent saves time by taking actions on your behalf. In exchange, you give it permissions and accept that mistakes will happen. The art is choosing which permissions, and which mistakes are recoverable.

---

**2 WHAT IS ALREADY IN REACH**

---

# Five automations that work today.

*Each one is something a UK SME has run in production in the last twelve months.*

These are not hypothetical. They are the most common agent builds we see working for small UK businesses. They are ordered from least to most consequential. Start with the first. They each describe what the agent does, what it needs, and what the human still does.

## 1. Monitor a LinkedIn profile and draft engagement comments

**What it does.** Watches a person's LinkedIn activity. Drafts a short, on-topic comment when they post. Sends the draft to the human for approval.

**What it needs.** The LinkedIn account of the person being monitored. The voice guidelines (from Edition 04). A draft queue in your inbox.

**What the human does.** Edits the draft if needed. Posts it. Replies to anything that warrants a reply.

## 2. Triage inbound email and draft replies

**What it does.** Reads new inbox messages. Sorts them by intent. Drafts a reply for the routine ones. Flags the rest.

**What it needs.** An inbox the agent can read, the voice guidelines, a list of intents and the right reply for each.

**What the human does.** Reviews the drafts. Sends the routine ones. Handles the flagged ones personally.

## 3. Watch a competitor's pricing page and alert on changes

**What it does.** Polls a list of competitor URLs on a schedule. Compares to the last version. Notifies a human when the content has changed.

**What it needs.** The URL list, a notification channel, the alert criteria.

**What the human does.** Reads the alert. Decides whether to respond. Updates the pricing brief.

## 4. Summarise weekly sales pipeline changes into a Monday briefing

**What it does.** Pulls pipeline data each Friday. Produces a short briefing with the changes, the deals at risk and the upcoming closes.

**What it needs.** Read access to the CRM, the output template, the delivery channel.

**What the human does.** Reads the briefing. Corrects the framing if it is wrong. Acts on the flagged items.

## 5. Scan incoming invoices against purchase orders and flag mismatches

**What it does.** Reads each new invoice. Matches it to the purchase order. Flags mismatches in price, quantity or supplier.

**What it needs.** Read access to the invoice inbox, access to the PO database, the alert channel.

**What the human does.** Reviews flagged invoices. Approves or queries them.

3 AN HONEST BREAKDOWN

# What an agent costs.

*API fees, setup time, ongoing care. And the math that decides whether it is worth it.*

The headline cost of an agent is small. The total cost, including setup and maintenance, is not. Be honest about both before you build.

## API costs

Most agents run on top of a model API. The price depends on the model and the volume of work. As a rough order of magnitude.

| WORKLOAD                                 | MONTHLY API COST | NOTES                                                            |
|------------------------------------------|------------------|------------------------------------------------------------------|
| Light (a few hundred requests per month) | £5 - £20         | Single small workflow, low concurrency.                          |
| Medium (a few thousand per month)        | £30 - £80        | Multiple workflows, daily runs, modest concurrency.              |
| Heavy (tens of thousands per month)      | £100 - £300+     | Hourly runs, multi-source, broader scope. Worth a private model. |

The above is API fees only. Most platforms charge a small additional subscription per month for the orchestration layer. That is usually £20 to £50 per month per agent.

## Setup time

Two to eight hours of focused work for a single agent that uses one or two external tools. Longer if you are integrating with systems that lack clean APIs. The first agent always takes longer than you expect. The second is faster. By the fifth you have a pattern.

## Ongoing maintenance

Thirty minutes per week is a realistic average for a working agent. Add more if the underlying systems change. Add more if the agent is doing consequential actions. Add less if the agent only watches and reports.

## The break-even calculation

$$\text{Monthly cost} = \text{API} + \text{subscription} + (\text{setup hours} \div 12) + (\text{weekly hours} \times 4.3) \times \text{hourly cost}$$
$$\text{Monthly benefit} = \text{human hours saved} \times \text{hourly cost} \times \text{realisation rate}$$

**HONESTY** If the break-even calculation only works when the agent is handling thousands of items a month, the agent is the wrong tool. Use a script, a SaaS subscription, or the human who was doing it already.

4 HOW MUCH ROPE THE AGENT GETS

# The permission ladder.

*Five rungs. Pick one. Most agents should start at rung two.*

Every agent has a permission level. The level sets what the agent can do without asking. Higher is not better. Higher means faster, more autonomous, more consequential. Lower means slower, more supervised, more reversible. Pick the lowest rung that gets the job done.

- 1

**Suggest.**  
The agent produces text or analysis. A human does everything that follows.No external actions. Suitable for research, summarisation, drafting.
- 2

**Draft.**  
The agent produces text ready to send. The human reviews and clicks send.Low risk. Most pilots should start here. Easy rollback is just deleting the draft.
- 3

**Stage.**  
The agent prepares the action and queues it. The human clicks confirm.Reversible at the click of a button. Appropriate for first contact with external systems.
- 4

**Execute with confirmation.**  
The agent executes within agreed limits. Out-of-bounds actions pause for human review.Faster, but requires the limits to be defined in writing. Suitable for routine, well-defined work.
- 5

**Execute within limits.**  
The agent operates within written limits and logs everything. Humans review by exception.Highest autonomy. Requires logging, monitoring and a way to stop the agent fast.

## Where the five automations sit

| AUTOMATION                | RUNG | WHY                                   |
|---------------------------|------|---------------------------------------|
| LinkedIn comment drafts   | 2    | Reputation risk. Human posts.         |
| Email triage and drafts   | 2    | Voice and relationships. Human sends. |
| Competitor pricing watch  | 1    | Watch and notify only.                |
| Monday pipeline briefing  | 1    | Briefing delivered. Human acts.       |
| Invoice mismatch flagging | 3    | Human approves each flagged invoice.  |

**5 HONEST LIST, WITH THE FIX FOR EACH**

# What can go wrong.

*Four ways agents fail. None of them are exotic. All of them are preventable.*

Agents fail in predictable ways. Knowing the categories in advance lets you build the right guardrail. Most agent incidents fit into one of the four below.

## Hallucinated actions

The agent invents a fact and acts on it. It files a ticket with the wrong customer number. It sends a reply that references a project that does not exist. The fix is the instructions from Edition 04. The agent should be told to flag, not guess, when the input is ambiguous. A budget cap and a human-in-the-loop on irreversible actions catch the rest.

## Runaway costs

The agent enters a loop, calls the model thousands of times, and the bill arrives on the first of the month. The fix is a hard cap on API spend per run, per day and per month. Most platforms let you set these. Set them lower than you think is necessary for the first month.

## Silent failures

The agent is running but the outputs are no longer useful. Nothing fails loudly. The human stops noticing. The fix is a weekly output sample, reviewed by the human. Five random outputs, every Friday. Ten minutes of attention catches drift long before it matters.

## The agent that emails the wrong person

The agent sends a draft intended for one contact to another. The fix is action whitelists, with a list of allowed recipients, allowed tools and allowed message templates. Anything outside the list pauses for human review.

**CONTAINMENT RULE** If an agent has been running for a month without incident, that is not evidence it cannot fail. It is evidence you have not yet seen the failure mode. Keep the logging on. Keep the weekly sample. Keep the kill switch easy.

## 6 THREE PATHS TO YOUR FIRST AGENT

# Build, buy, or hire.

*Three paths. The right one depends on the task and the team.*

There are three ways to get an agent working in your business. Build it yourself with no-code tools. Buy a vertical SaaS that does one thing well. Hire someone to build and maintain it for you. The right choice depends on how standard the task is, how unique your business is, and how much time you have.

### BUILD

**Cost.** Your time. Two to eight hours per agent for the first one.

**Tools.** Hermes, n8n, Make, Zapier or similar.

**Best for.** Unique workflows. Tasks that need access to your specific systems.

**Trade.** You own the work and the maintenance.

### BUY

**Cost.** £20 to £200 per month per tool.

**Tools.** Vertical SaaS that does the one thing you need.

**Best for.** Standard tasks. Lead enrichment, scheduling, transcription.

**Trade.** You do not own the workflow. The vendor can change it.

### HIRE

**Cost.** £500 to £2,000 for setup. £200 to £500 per month for maintenance.

**Best for.** Mission-critical workflows. Anything you cannot afford to be wrong.

**Trade.** Fastest path to a working system, but the cost only works for high-value workflows.

## Decision framework

Three questions, in order. The first question whose answer is "yes" picks the path.

1. **Is the task standard, and does an off-the-shelf tool cover it?** Buy.
2. **Is the task unique to your business, and do you have the time to learn the tools?** Build.
3. **Is the task unique, valuable, and time-critical?** Hire.

**Default for most UK SMEs.** Start with "build" for the first agent. The first one teaches you what the others will cost. Skip to "hire" when you can describe the third agent clearly in writing. Skip to "buy" when an off-the-shelf tool demonstrably covers the task.

7 PICK THE SIMPLEST ONE AND SHIP IT

# Your first agent, in two hours.

*The competitor pricing watch. The least risky of the five.*

If this is your first agent, pick the competitor pricing watch from page 3. It is read-only. The agent watches, drafts an alert, and notifies a human. There is no outbound action to reverse. Two hours of focused work is enough to get a working version.

## What to connect

- A list of competitor URLs. Three is a sensible minimum, ten a practical maximum for the first version.
- A notification channel. Email or Slack, whichever your team actually reads.
- An output destination for the captured content. A shared folder is fine.

## What to configure

- A polling schedule. Hourly is fine for a price page. Daily for everything else.
- A change threshold. Anything above a small percentage, or any change to a specific field.
- The alert format. Three lines: what changed, the previous value, the new value, plus a link.
- A quiet hours rule. No alerts between 10pm and 8am unless the change is over the threshold.

## What to test

Before you let it run unattended, change something on one of the competitor pages yourself. Confirm the alert fires, the captured content is right, and the link works. Do this twice. Then turn it on.

## What to monitor in the first week

| DAY | CHECK                                               |
|-----|-----------------------------------------------------|
| Mon | Three alerts reviewed. Any false positives?         |
| Tue | Threshold tested. Adjust if too noisy or too quiet. |
| Wed | Notification channel. Did it actually get read?     |
| Thu | Captured content folder. Is the format usable?      |
| Fri | Cost so far. Within budget?                         |

### END OF WEEK ONE

Decide whether to keep the threshold, adjust it, or add a second watch. The first agent is rarely the last. Treat it as a working prototype, not a finished product.

## 8 THREE SIGNS AN AGENT IS THE WRONG TOOL

# When to stop.

*Not every task should be automated.*

It is easy to lose a week to a clever agent build for a task that should never have been automated. The pull is strong. The capability is new. The instinct to apply it to anything structured is human. Resist it for three categories of task.

### 1. The task requires genuine judgement

An agent can draft a difficult customer reply. It should not be the one sending it. An agent can summarise a complex negotiation. It should not be the one deciding the terms. Where the answer depends on weighing factors that humans understand but cannot easily list, the human stays in the loop.

### 2. The error cost is catastrophic

Sending the wrong invoice. Filing a regulatory document with the wrong data. Terminating a contract by mistake. If a single mistake could cost more than the entire saving from automation, the task is not a candidate. Do the work manually. Slow is fine when fast and wrong is not.

### 3. The volume is too low to justify setup

An agent that handles twenty items a month, at eight hours of setup, is not a sensible automation. The break-even calculation from page 4 is honest about this. Two hours of human work a month is below the threshold for almost every agent build.

## What to do instead

- **For judgement-led tasks.** Use the agent to draft, the human to decide. Edition 04's six-part instruction covers the structure.
- **For high-stakes tasks.** Do them by hand, and write down the steps. A well-written checklist is often the right answer.
- **For low-volume tasks.** Leave them. Revisit when the volume changes, or when the human gets tired of them.

**Honest test.** If a colleague heard you describe the agent and asked "why would you automate that?", you have your answer. Stop the build. Spend the time on the task that the colleague suggested you were avoiding.

---

**9 IF YOU WANT ONE RUNNING BY NEXT MONDAY**

---

# How Expedient AI can help.

*Plain words about what we do and what it costs.*

Expedient AI builds and maintains AI agents for UK small and medium-sized businesses. We started in 2024 in Bristol. The team is small, the work is hands-on, and most of our clients find us through word of mouth rather than through marketing.

## What we do

Three things. We scope the first agent with you, in plain language, against a real workflow. We build it on the platform that fits the task, with the permissions and guardrails described in this guide. We maintain it, with a named person who knows your business, monthly reviews and a 24-hour response on incidents.

## What that costs

A first agent, scoped and built, runs from £1,500 to £4,000 depending on the integration work. Maintenance is £250 to £500 per month, which covers a monthly review, the named person, and time budgeted for incidents and small changes. Larger systems cost more. We will tell you the price after the scoping call, not before.

## What it looks like in practice

A typical first engagement lasts four weeks. Week one is the scoping. We talk to the people who do the work, write the pilot canvas from Edition 01, and agree the first claim. Week two is the build, against a closed test environment. Week three is the offline test from Edition 02, days 6 to 10. Week four is the limited live pilot, with a written decision meeting at the end.

## What we will not do

- Sign a multi-year contract before the first agent has run for sixty days.
- Take on a workflow where the data tier is Restricted without a documented basis.
- Tell you the agent will replace a member of staff.
- Charge for a meeting to decide whether we should meet.

## How to start

Email with one paragraph on the workflow you want to automate. We will reply within two working days with either a scoping proposal or a polite suggestion that this guide is enough for now.

Website: [expedientai.co.uk](https://expedientai.co.uk)

**10 WHAT TENDS TO RECUR**

# Lessons from the field.

*Patterns observed across agent builds in 2025 and 2026.*

These notes come from agent builds run with UK small and medium-sized businesses over the last eighteen months. They are observations, not rules. Read them, decide what applies to your situation, and make your own call.

## What tends to work

- Starting with the smallest agent that demonstrates value. The first win funds the second.
- Picking a workflow the human is tired of. Tired humans write the best prompts.
- Keeping the human in the loop for the first month, even on routine actions.
- Writing the kill switch into the system before turning the system on.
- Reviewing the output sample every Friday, even when nothing has gone wrong.

## What tends to fail

- Automating a workflow that is not yet stable. Fix the process, then automate.
- Granting the agent the same permissions as the human, on day one.
- Skipping the data classification from Edition 03 because the workflow is "internal".
- Letting an agent send external messages without a recipient whitelist.
- Treating a month of working output as proof the agent cannot fail.

## Three questions to ask before each build

1. If this agent is wrong once a month, what is the worst that happens?
2. If the vendor's API goes down for a day, what is the visible impact?
3. If the human leaves, can someone else pick up the agent within a week?

If the answer to any of these is unmanageable, do not build the agent as designed. Adjust the design, or pick a different workflow. The whole point of starting small is that you can answer these questions cheaply.

**One last thing.** If you build one agent and it works, you will be tempted to build ten. Build three. The fourth is where most teams run out of maintenance capacity.

**11** COLOPHON

# The series and how to reach us.

*Six field guides for UK SMEs. One Bristol-based company behind them.*

## The series

Six field guides for UK small and medium-sized businesses. Independent, plain-English, written by a practitioner.

**01 · Where AI Actually Helps**

Choose one worthwhile use case, scope a 30-day pilot and estimate the return before buying software.

**02 · The 30-Day AI Pilot**

A day-by-day guide to running the pilot you scoped in edition 01.

**03 · Your Data, Your Rules**

AI and privacy in plain English for UK businesses handling personal data.

**04 · Writing AI Instructions That Actually Work**

The six-part structure that turns vague prompts into reliable outputs.

**05 · The AI Vendor Interrogation Kit**

Eight questions to ask before signing a contract with any AI vendor.

**06 · Your First AI Employee**

A non-technical guide to autonomous AI agents for small businesses.

## Colophon

**PUBLICATION**

Your First AI Employee

July 2026 · Version 1.0

Author: Ted Milton, Expedient AI Ltd

**CONTACT**

Expedient AI Ltd

[expedientai.co.uk](https://expedientai.co.uk)

You can use this guide without contacting Expedient AI. If you would value an independent review of your data handling or AI controls, Expedient AI Ltd offers practical AI implementation support for UK SMEs.