

A PLAYBOOK FOR UK BOARDS AND EXECUTIVES

The AI Control Playbook

How to see, cut, contain, prove and scale the AI already
inside your business.

HOW TO USE THIS PLAYBOOK

Read the summary on page 3. If you only act on one thing, act on page 16: the register. Everything else in this document depends on it.

The playbook is written for the people who sign off AI spend and answer for AI risk: chief executives, finance directors, chief operating officers, general counsel and the board. It assumes no technical knowledge. Every figure is sourced on page 24, and vendor-commissioned surveys are marked as such.

Pages 16 to 21 are templates. Print them. Put an owner’s name against each line. Re-run the self-check on page 21 at the end of the 90 days and compare.

Second edition, September 2026. The first edition, The AI Visibility Playbook, covered seeing the estate. This edition adds the four moves that follow: cut, contain, prove and scale. It reflects the model-pacing announcements of September 2026 and the UK regulatory changes in force since February 2026.

CONTENTS

SUMMARY	Every business now runs on AI. Few are in control of it.	3
CONTEXT	What is actually happening in UK businesses	4
DIAGNOSIS	One problem wearing six hats	5
CASE NOTES	A breach, and a budget that ran out	6
PROBLEMS	01 The meter you can't see	8
	02 Premium by default	9
	03 Accounts you don't own	10
	04 The same work twice	11
	05 Skill trapped in people	12
	06 The reckoning	13
OUTSIDE	The two underneath: compliance and turbulence	14
METHOD	Five moves. One loop.	15
SEE	The AI register	16
CUT	The savings you are not taking	17
CONTAIN	Rules as controls, through one gateway	18
PROVE	The one-page board report	19
SCALE	The 90-day plan	20
CHECK	The seven-question self-check	21
TAKE-AWAYS	Seven things to take into the next board meeting	22
GENWA	How GenWA helps	23
	Sources	24

SUMMARY

Every business now runs on AI. Few are in control of it.

AI arrived in most UK businesses through the side door: personal accounts, expense claims and features switched on inside software you already owned. Nobody instrumented it. The bill is now metered by use rather than by seat, so it grows without anyone deciding that it should. And the rules that hold a named human accountable for the output are already in force.

This playbook sets out six problems we see in almost every UK mid-market business, each with a sourced figure, what it costs and one thing to do this week. Underneath the six sit two more: compliance, where the law changed in February 2026, and turbulence, where the models your people use change every few months and the choice is left to whoever is at the keyboard.

They are one problem wearing six hats. Fear-driven adoption produced a sprawl of tools and accounts on a pricing model that compounds, used by people who do not share what works, under rules that now demand a record. The answer is not to spend less on AI. It is to put the board back in control of it.

FIVE MOVES, IN ORDER

See every tool, account, agent and pound, with an owner. **Cut** the cost by routing each task to the cheapest adequate model. **Contain** the risk by enforcing the rules on every request, not on the intranet. **Prove** it with a record an auditor, insurer or regulator can read. Then **Scale**: deploy agents on a foundation that can hold them.

71%

of UK employees have used AI tools their employer never approved.

Microsoft UK, 2025 · vendor survey

56%

of AI cost now goes to premium models. Ten months earlier it was 6%.

Ramp, April 2026 · vendor data

4×

Businesses with full visibility of AI cost report an established return at four times the rate of those without: 25% against 6%.

KPMG UK, July 2026

WHY BOTHER

The businesses that win with AI are not the ones spending most. They are the ones that can see it, and can show what it did.

CONTEXT

What is actually happening in UK businesses

Eight facts, official sources first. Adoption has roughly tripled in three years. Policy, training and control have not kept up.

35%

of UK businesses with 10+ staff now use at least one AI technology, up from about 12% in September 2023. Among 250+ staff it is 49%; only 10% use AI extensively.

ONS, July 2026

54%

of UK SMEs are actively using AI, up from 35% a year earlier.

British Chambers of Commerce, March 2026

76%

of UK businesses rate high cost as a significant barrier to AI; 72% say unclear regulation; 60% limited skills.

DSIT, February 2026

24%

of UK businesses using or considering AI have cyber-security practices that address AI risk.

DSIT Cyber Security Breaches Survey, April 2026

46%

of UK financial services firms have only a partial understanding of the AI they use. 75% use AI.

Bank of England and FCA, 2024

31%

of UK employers have a formal policy on generative AI. 35% provide training. 25% prohibit it.

CIPD, Autumn 2025

71%

of UK employees have used unapproved consumer AI at work; 51% weekly; 28% say no approved alternative exists.

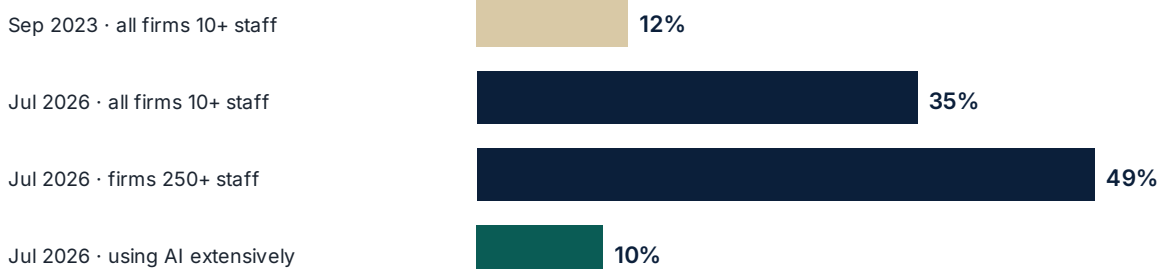
Microsoft UK / Censuswide, 2025 · vendor survey

42%

of UK leaders have only partial visibility of AI spend; 33% do not understand AI cost structures, tokens included.

KPMG UK, July 2026

ADOPTION HAS TRIPLED. EXTENSIVE USE HAS NOT.



Share of UK businesses using at least one AI technology. ONS, Artificial intelligence in UK businesses, July 2026.

READING THE ADOPTION NUMBERS

Official statistics (ONS, DSIT) put firm-level adoption at 16% to 35% depending on the size threshold and definition; business-bank and chamber surveys, which count any use among engaged firms, put it at 54% to 66%. For a business of 50 to 2,000 staff the defensible line is that roughly a third to a half use AI, and about one in ten use it extensively. What no survey counts is the AI that walked in on a personal account, which is the subject of page 10.

DIAGNOSIS

One problem wearing six hats

Each of the six problems on the next pages has its own figure, its own cost and its own fix. But they share one cause: the business cannot see what it is using, so it cannot decide, control or measure it.

The meter you can't see 01

AI is billed by use, use is compounding, and only 26% of companies fully track the cost.

Accounts you don't own 03

One in three employees use a personal account for work AI. The business cannot see it or switch it off.

Skill trapped in people 05

Gains are real for individuals and invisible to the organisation. 81% of UK workers will not disclose their AI use.

Premium by default 02

Which model does the work is decided by a default setting, not by the task. Premium models take 56% of cost.

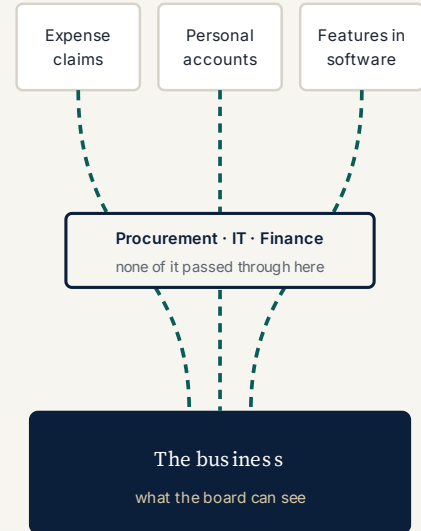
The same work twice 04

Teams buy, build and prompt the same capability separately. Half of deployed agents run in silos.

The reckoning 06

Investment made from fear now meets a CFO asking for the return. 42% of companies abandoned most AI initiatives in 2025.

HOW AI GOT IN



Business units, not IT, control 81% of software spend (Zylo, 2026, vendor data). ChatGPT is now the single most-expensed application in business (Zylo, 2026).

CASE NOTE · COMPLIANCE

Three engineers, twenty days, and the source code was gone

20 days

from permitting a consumer AI tool to the first leak

3

separate incidents: source code, test code, meeting minutes

£0

fine. The cost was trade secrets, and a company-wide ban

WHAT HAPPENED

In March 2023 Samsung's semiconductor division allowed engineers to use ChatGPT. Within about twenty days three of them had pasted confidential material into the public service: proprietary source code, to find a bug; the code for a chip test sequence, to optimise it; and the transcript of an internal meeting, to generate minutes. At the time the service used conversations for training by default and offered no straightforward way to delete them. In May Samsung banned generative AI tools on company devices and networks and told staff not to enter "any personal or company related information" into them (CNBC, May 2023).

Nobody broke a rule, because there was no rule. Nobody bypassed a control, because there was none. The tool was the easiest thing available and the work was real. That is how most AI enters a business, and it is why the UK's solicitors' regulator now says that putting client material into an open tool "is to place this information on the internet in the public domain" (SRA warning notice, August 2026).

WHY IT MATTERS HERE

The same pattern, one level up, is the reason a Canadian tribunal held Air Canada liable for what its chatbot told a customer (Moffatt v Air Canada, February 2024: "it should be obvious that it is responsible for all the information on its website"), and why the UK High Court referred a barrister and a solicitor to their regulators for filing AI-invented case law (Ayinde v Haringey, June 2025). The organisation owns the output. Regulators, courts and insurers are aligned on that.

THE CONTROL THAT WOULD HAVE CAUGHT IT

See	A register would have shown the tool in use on day one, on which accounts, by whom.
------------	---

Contain	A gateway detects source code, client material and personal data in a request and blocks or redacts it before it leaves. The rule exists as a control, not a memo.
----------------	--

Prove	The record shows what was tried, what was stopped and why. That is the evidence the ICO, an insurer or a customer asks for.
--------------	---

Sources. CNBC, 2 May 2023; Forbes, 2 May 2023; SRA, Misuse of AI warning notice, 17 August 2026; Moffatt v Air Canada, 2024 BCCRT 149; Ayinde v London Borough of Haringey [2025] EWHC 1383 (Admin). These are public incidents, reported by others; GenWA had no involvement in any of them.

CASE NOTE · OVERSPEND

The budget for the year lasted four months

4 months

to spend a full year's AI tooling budget

\$1,500

the monthly cap per employee, per tool, imposed afterwards

“as much as possible”

the instruction staff had been given

WHAT HAPPENED

Uber told its engineers to use AI coding tools as much as possible and ran internal leaderboards to encourage it. The tools were priced by consumption. By April 2026 the company had spent its entire annual budget for them. In June it capped every employee at \$1,500 a month per agentic coding tool, added dashboards and an approval route for exceptions (Bloomberg, 2 June 2026; TechCrunch, 2 June 2026). Its chief operating officer said it was “very hard to draw a line” from AI usage to new customer features.

Nothing went wrong with the tools. The tokens bought real work. What was missing was a meter anyone could see, a limit anyone had set, and a measure of return anyone had agreed. Consumption pricing with encouragement and no control is a budget with no bottom.

WHY IT MATTERS HERE

Uber is unusual only in being reported. Across Ramp’s customer base token usage rose 1,001% in fifteen months while spend rose 497% (Ramp, April 2026); 62% of organisations have had unexpected AI costs large enough to change a business decision (Mavvrik, 2026, vendor survey); and when Cursor moved its plans to consumption pricing in June 2025, users burned through allowances in days and the company apologised and refunded (TechCrunch, July 2025). The pattern is structural, not a Silicon Valley story.

THE CONTROL THAT WOULD HAVE CAUGHT IT

See A live meter by team, tool and person, with the month’s total against budget, visible before the invoice.

Cut Routing sends routine requests to the cheapest adequate model. Most of a coding tool’s tokens are routine. Per-person and per-agent caps pause work at the limit rather than bill over it.

Prove Three use cases under measurement, each with a number, a review date and a stop criterion, so “hard to draw a line” becomes a line.

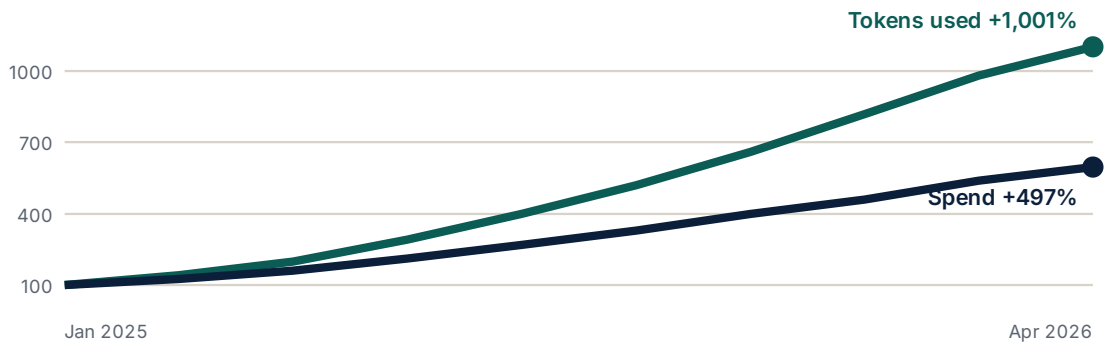
Sources. Bloomberg, 2 June 2026; TechCrunch, 2 June 2026; Ramp AI Index, April 2026 (vendor data); Mavvrik, 2026 (vendor survey); TechCrunch, 7 July 2025. These are public incidents, reported by others; GenWA had no involvement in any of them.

01

The meter you can't see

PROBLEM 1 OF 6

FIXED BY · SEE · CUT



Indexed, January 2025 = 100. Token usage across Ramp's customer base rose 1,001% in fifteen months while spend rose 497%. Ramp, April 2026 (vendor data).

THE PROBLEM IN PLAIN ENGLISH

Until recently AI was bought like software: a known number of users at a known price. That has gone. Most AI is now billed by consumption, and consumption is compounding. Unit prices are falling fast, by 9 to 900 times a year for a fixed level of capability (Epoch AI, 2025), and bills are rising anyway, because use grows faster than prices fall.

Almost nobody can see it happening. Only 26% of companies fully track AI costs and 22% see nothing until the invoice arrives (KPMG, 2026). Only 11% can forecast AI spend to within 10% (Mavvrik, 2026, vendor survey). The cause is structural: the people who see usage have no budget authority, and the people who pay see the bill weeks later.

WHAT IT COSTS

Overruns discovered after the fact. 62% of organisations have had unexpected AI costs large enough to change a business decision; a third imposed emergency spending freezes (Mavvrik, 2026, vendor survey). Uber exhausted its entire 2026 AI budget in four months after telling staff to use AI as much as possible (Bloomberg, 2026). CIOs can misjudge AI scaling costs by 500% to 1,000% (Gartner, 2024).

26%

of companies fully track what their AI costs.
22% see nothing until the invoice arrives.

KPMG, 2026

11%

of enterprises can forecast AI spend to within 10%.

Mavvrik, 2026 · vendor survey

REMEMBER

×10

token growth in fifteen months

ASK YOUR TEAM

What did we spend on AI last month, to the pound, and who signed it off?

DO THIS WEEK

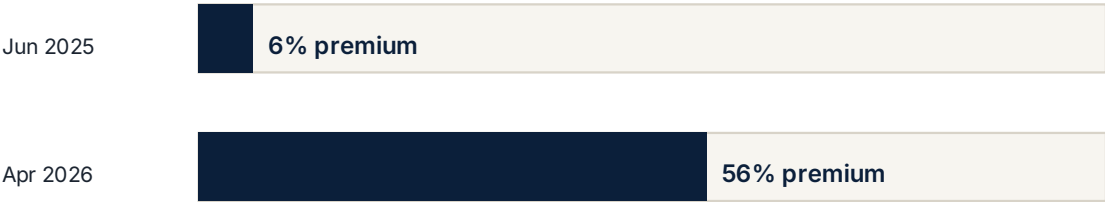
Ask finance for every AI-related line in the last three months of expense claims, card statements, subscriptions and cloud bills. Put them in one list with the team that incurred each. Only 26% of companies fully track AI costs, so expect the total to surprise you. You now have your first month's figure, and the baseline for the register on page 16.

02

Premium by default

PROBLEM 2 OF 6

FIXED BY · CUT



Premium models' share of total AI cost. They account for 46% of tokens but 56% of cost. Ramp, April 2026 (vendor data).

THE PROBLEM IN PLAIN ENGLISH

Inside most businesses, which AI model gets used for a task is decided by whatever the tool defaulted to, not by what the task needs. The price spread between the cheapest and most expensive mainstream models is about 20 times, for tasks on which, in the words of one survey of 100 enterprise CIOs, “all models perform well enough now” (a16z, 2025).

The vendors say the same. Anthropic’s own documentation states that unexpectedly high spend “usually traces back to long sessions that were never cleared or to Opus left as the default model” (Anthropic, 2026). When OpenAI switched on automatic routing, daily use of its expensive reasoning models among paying users jumped from 7% to 24% overnight (OpenAI, 2025). That is how little of model choice is a considered decision.

WHAT IT COSTS

Frontier prices for routine work. Routing tasks to the cheapest adequate model retains about 95% of frontier quality at up to 85% lower cost (LMSYS RouteLLM, 2024). Sprawl compounds it: companies using 26 or more models spend \$442 per employee per month on tokens against \$28 for those using four to ten (Ramp, 2026, vendor data). Page 17 shows what this is worth to a 300-person firm.

85%

lower cost at 95% of frontier quality when tasks go to the cheapest adequate model.

LMSYS RouteLLM, 2024

20×

price spread between the cheapest and most expensive mainstream models.

a16z, 2025

REMEMBER

56%

of cost goes to premium models

ASK YOUR TEAM

Which model does our routine work, and who decided that?

DO THIS WEEK

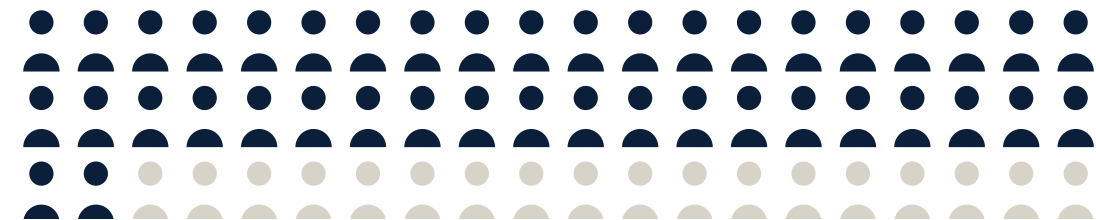
Open every AI tool the company pays for and check its default model. Set the default to the standard tier and make the premium tier a deliberate choice, with a one-line note on when it is worth it. Do the same in any internal code that calls a model API. This takes an afternoon and is usually the single largest saving available.

03

Accounts you don't own

PROBLEM 3 OF 6

FIXED BY · SEE · CONTAIN



Of every 60 UK employees, 42 have used unapproved consumer AI tools at work. Microsoft UK / Censuswide, 2025 (vendor survey).

THE PROBLEM IN PLAIN ENGLISH

A large share of the AI used for work runs on accounts the business does not own, cannot see and cannot switch off. Roughly one in three employees access AI through a personal account; for some tools it is a majority (Cyberhaven, 2026, vendor telemetry). 90% of employees use personal AI tools for work while only 40% of companies have an official subscription (MIT NANDA, 2025). In the UK, 28% of employees say there is no approved alternative (Microsoft UK, 2025, vendor survey).

The spend follows the same route. Expense-report software spend is up 267% year on year (Zylo, 2026, vendor data). It arrives through expenses, not procurement, so nobody whose job it is to control it ever sees it.

WHAT IT COSTS

Data leaving through channels you do not control. Roughly a third of the data entering AI tools is sensitive (Cyberhaven, 2025, vendor telemetry). 63% of UK organisations lack AI access controls, and where shadow AI use is high the cost of a breach rises by £498,000 against a UK average of £3.78 million (IBM, 2025, vendor research). 92% of organisations that suffered an AI-related breach in 2026 had no AI-specific access controls (IBM, 2026).

£498k

added to the cost of a UK data breach where shadow AI use is high, against an average of £3.78m.

IBM, 2025 · vendor research

1 in 3

employees use a personal account, not a company one, for AI at work.

Cyberhaven, 2026 · vendor telemetry

REMEMBER

71%

of UK staff have used unapproved AI

ASK YOUR TEAM

If we switched off every personal account tomorrow, what work would stop?

DO THIS WEEK

Give every employee one approved AI tool on a company-owned account and tell them so in writing. Say what it may be used for and what must never go into it. A quarter of unapproved use exists because there was no approved alternative. Then ask people to move their work across; do not ban the personal account yet, or you will lose sight of it entirely.

04

The same work twice

PROBLEM 4 OF 6

FIXED BY · SEE · CUT

Typical organisation



Top 1% of organisations



Generative AI applications in use. The typical organisation runs six to eight, up 33% in a year; the top 1% run between 47 and 89 (midpoint shown). Netskope, 2026 (vendor data).

THE PROBLEM IN PLAIN ENGLISH

Teams are buying, building and prompting the same capability separately, without knowing it. Generative AI entered the list of the most redundant software functions for the first time in 2026 (Zylo, 2026). Half of deployed AI agents operate in isolated silos, and only 54% of organisations have any central oversight of them (Salesforce/MuleSoft, 2026, vendor survey). Only 5% of internally built AI tools ever reach production (MIT NANDA, 2025).

Duplication also shows up as rework. 40% of desk workers received “workslop” in the past month: AI output that looks finished and has to be redone, at about two hours per instance (BetterUp and Stanford, in Harvard Business Review, 2025).

WHAT IT COSTS

Overlapping subscriptions, rebuilt automations and rework. 36% of software licences go unused (Zylo, 2026, vendor data). Workslop costs about \$186 per employee per month (BetterUp and Stanford, 2025). 77% of employees say AI has increased their workload and 39% now spend more time reviewing AI output (Upwork, 2024, vendor survey). There is no published figure for how often the same prompt or agent is built twice inside one business; we expect it to be high.

36%

of software licences go unused.

Zylo, 2026 · vendor data

50%

of deployed AI agents operate in isolated silos.

Salesforce / MuleSoft, 2026 · vendor survey

REMEMBER

6–8

AI apps in the typical firm

ASK YOUR TEAM

How many of our AI tools do the same job as another one we already pay for?

DO THIS WEEK

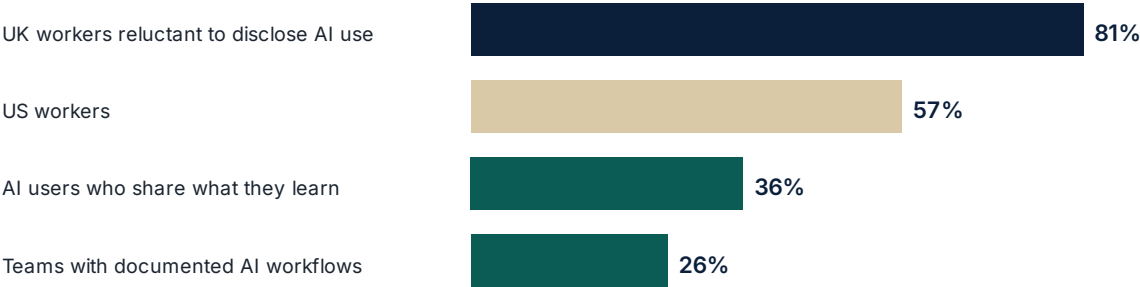
Take the list from page 8 and sort it by department. Circle every pair of tools that does the same job. Then ask each department head one question: “What have you built or set up in AI that another team might already have?” Put the answers in a shared place with a name against each. That is the start of your register.

05

Skill trapped in people

PROBLEM 5 OF 6

FIXED BY · SEE · SCALE



Cornerstone OnDemand, 2025 (vendor survey); Microsoft Work Trend Index, 2026 (vendor survey).

THE PROBLEM IN PLAIN ENGLISH

The gains from AI are real for individuals and largely invisible to the organisation. 80% of people report personal productivity gains; only 37% of organisations report any effect on profit, a figure that has not moved in a year (McKinsey, 2026). The reason is that AI use is private. 57% of employees hide their AI use and present the output as their own (KPMG and University of Melbourne, 2025).

The UK is worse than most. 51% of UK workers never or rarely receive AI training (Cornerstone, 2025, vendor survey). Only 11% of UK businesses say more than half their workforce has been trained (ONS, 2026). Only 35% of UK employers provide any AI training (CIPD, 2025).

WHAT IT COSTS

You pay for AI at organisational scale and receive returns at individual scale. Every working prompt and lesson learned is rediscovered by isolated people who then change roles or leave. Organisational factors (culture, manager support, documented workflows) explain twice as much of AI's impact as individual skill: 67% against 32% (Microsoft, 2026, vendor survey). Because use is hidden, the business cannot even see what it is losing.

81%

of UK workers are reluctant to disclose their use of AI at work.

Cornerstone OnDemand, 2025 · vendor survey

37%

of organisations report any effect of AI on profit, unchanged in a year.

McKinsey, 2026

REMEMBER

36%

of AI users share what they learn

ASK YOUR TEAM

Who in this business is best at AI, and where is what they know written down?

DO THIS WEEK

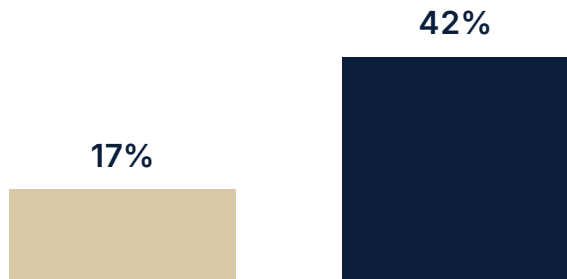
Ask each team lead for the three AI prompts, workflows or tools their team actually uses and finds valuable. Put them in one shared folder, with the author's name and a line on what it is for. Then say, in public, that using AI well is something the business wants to hear about. In a country where four in five workers will not admit to using AI, that sentence from the top matters more than any tool.

06

The reckoning

PROBLEM 6 OF 6

FIXED BY · PROVE



Share of companies that abandoned most of their AI initiatives. S&P Global, 2025.

THE PROBLEM IN PLAIN ENGLISH

A large share of AI investment was made without a business case. 64% of CEOs admit that “the risk of falling behind drives investment in some technologies before they have a clear understanding of the value” (IBM, 2025). 58% of UK IT leaders say they invested in AI out of fear of being left behind (ABBYY, 2024, vendor survey). 75% of executives admit their AI strategy is “more for show” (WRITER, 2026, vendor survey).

The pressure is now reversing. 56% of CEOs report no significant financial benefit from AI so far (PwC, 2026). 49% of leaders have scaled back or paused agent rollouts because operating costs outpaced value (KPMG, 2026). Gartner predicts more than 40% of agentic AI projects will be cancelled by the end of 2027 (Gartner, 2025).

WHAT IT COSTS

Sunk cost, cancelled projects and a credibility problem for whoever sponsored the spend. The CFO is now asking questions the CIO cannot answer, and the CEO who championed AI in 2024 needs a defensible story in 2026. 92% of CFOs feel pressure to demonstrate a return on AI (Avalara, 2026, vendor survey). 82% of senior leaders are concerned about token costs and 98% of those using token-based tools say cost has made them reconsider their approach (EY, 2026).

Share of companies that abandoned most of their AI initiatives

56%

of CEOs report no significant financial benefit from AI to date.

PwC Global CEO Survey, January 2026

49%

of leaders have scaled back or paused agent rollouts because costs outpaced value.

KPMG Global AI Pulse, 2026

REMEMBER

42%

abandoned most AI initiatives in 2025

ASK YOUR TEAM

For each AI spend line: what does it do, for whom, and how would we know it was working?

DO THIS WEEK

For every AI spend line on your list, write the business case in one sentence: what it does, for whom, and how you would know it was working. If a line has no sentence, flag it. Do not cancel anything yet. The point is to know which spend is defended by evidence and which by fear, before the board asks. Page 19 shows how to report the answer.

OUTSIDE THE ESTATE

The two underneath: compliance and turbulence

The six problems are about seeing your own estate. Two more come from outside it: the rules have already changed, and the models change every few months.



COMPLIANCE: THE RULES ARE ALREADY IN FORCE

UK GDPR and the Data (Use and Access) Act

2025. The Act's main provisions commenced on 5 February 2026. It widens the scope for automated decision-making but attaches mandatory safeguards: information, a right to contest, and meaningful human intervention. Any business using AI on personal data is already inside these rules.

The ICO's statutory AI code. A statutory Code of Practice on AI and automated decision-making is required by SI 2026/425, in force from 12 May 2026, with final guidance due in winter 2026. Enforcement is real: the ICO's largest 2025 fine was £14 million, against Capita.

ISO/IEC 42001. The standard for AI management systems is becoming the buyer's question in tenders and insurance renewals. About 350 organisations worldwide were certified by April 2026; BSI became the first UKAS-accredited certifier in January 2026.

The gap. Only 31% of UK organisations have AI governance policies and 63% lack AI access controls (IBM, 2025). For AI agents, the NCSC asks for scoped credentials, logging, human oversight and a kill-switch (NCSC, 2026). If the regulator, your insurer or a customer asked tomorrow, what would you show them?

TURBULENCE: WHICH MODEL, FOR WHAT, AND WHEN NOT

The model market is being reshaped, at speed, by a handful of companies whose decisions you neither control nor see coming. New models arrive monthly. Older ones are retired within months, often with weeks of notice, and behaviour can change without a version number changing. Vendors alter prices and usage limits unilaterally.

On 14 September 2026 the chief executives of Anthropic, OpenAI, Google DeepMind and xAI backed a call to pace frontier development. Pacing applies to the next capability jump, not to the release cadence of the models your people already use. The median business already uses nine different models (Ramp, 2026). A policy written around one vendor's model is already out of date.

Regulators treat a model change as a governed event: the Bank of England's model risk standard covers it, the EU AI Act treats substantial modification as a new obligation, and the NCSC's guidance includes update management. None of that reaches an employee choosing between browser tabs.

WHAT GOOD LOOKS LIKE

An approved model list with a use policy per task: which model for which job, on which data, with which guardrails, and where AI must not be used. Set centrally, applied as the default, re-checked when a model changes. Decided once, for the business, not by each employee.

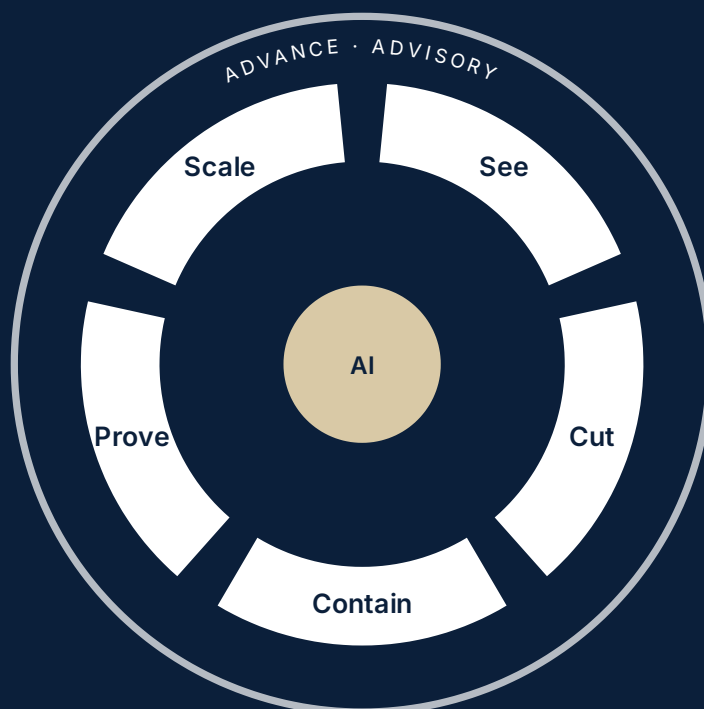
METHOD

Five moves. One loop.

The six problems are fixed in one order, and it is not optional. You cannot cut what you cannot see. You cannot prove what you did not contain. And you should not scale what you cannot prove.

Each move is a product family at GenWA and a page in this playbook. **See** gives you the register and the live picture of every tool, account, agent and pound. **Cut** routes every task to the cheapest adequate model and retires the seats and duplicates nobody uses. **Contain** puts the rules in the path of every request through one gateway, so a policy becomes a control. **Prove** generates the evidence from the record rather than assembling it the week before the visit. **Scale** deploys agents, individually or as teams, on a foundation that can hold them.

Around the five sits **Advance**: the advisory work of audit, policy, training and managed operations, for businesses that want the outcome without building the capability themselves.



AI at the centre, always moving. Five families hold it.
Advisory around it.

The labs have concluded that capability without control is reckless. That is a reasonable standard for a board to adopt.

01

See

Register every tool, account, agent and pound



02

Cut

Route each task to the cheapest adequate model



03

Contain

Enforce the rules on every request



04

Prove

Generate the evidence from the record



05

Scale

Deploy agents on a controlled foundation

The AI register

One line per AI asset: a tool, a subscription, an internal build, an agent, an automation, or an AI feature that arrived inside software you already pay for. This is the document every other control depends on.

ASSET	OWNER	PURPOSE	DATA IT TOUCHES	ACCOUNTS	MONTHLY COST	MODEL DEFAULT	STATUS
<i>Illustration:</i> ChatGPT (team plan)	Head of Marketing	Drafting copy, summarising research	Campaign briefs; occasional customer names	6 company, 4 personal	£240	Premium	Under review: move personal to company; set default to standard

HOW TO FILL IT IN

Asset. Name and plan, or the internal name for something you built. Include AI features inside existing software: 77% of IT leaders have found AI features arriving without their knowledge (Zylo, 2026, vendor data).

Owner. A named person, not a department. If nobody will own it, that is a finding.

Purpose. The job it does, in plain words. If you cannot write one, see page 13.

Data it touches. What goes in. Flag personal data, customer data, client material, commercial IP and anything special-category.

Accounts. How many company-owned, how many personal. The second number is your exposure.

Monthly cost. Subscription plus usage plus anything expensed. Write £0 if free; free tools still touch your data.

Model default. Which tier the tool uses unless someone changes it. This column is where page 17 starts.

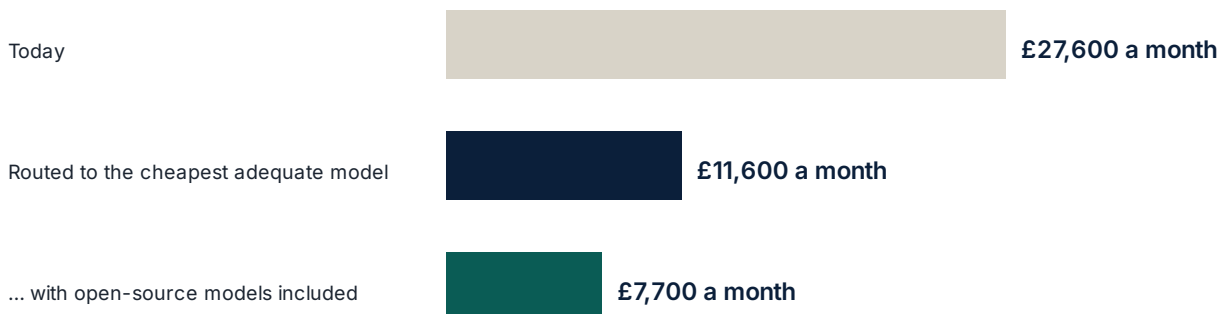
Status. Approved · Under review · Retire · Replace with (name).

With GenWA the register is not typed. Discover builds it from your usage data in 48 hours, keeps it current, and flags every new tool, account and feature as it appears.

MOVE 02 · CUT · WORKED EXAMPLE

The savings you are not taking

A 300-person business, with AI spend at the level our audits typically find. Nothing here requires anyone to use AI less.



Monthly AI spend at £92 per employee (Ramp, 2026: AI at 15% of software spend; Zylo, 2026: £7,400 software spend per employee). Standard routing: 56% of spend on routable tasks × 75% routed × 85% saving, plus 36% unused seats on 40% of spend, plus 8% duplicates. Open-source models: 85% routable at 92% cheaper plus 45% of the standard tier at 80% cheaper. Capped at 72%.

WHERE THE MONEY IS

Routing. The largest line. Most requests are routine: summarising, drafting, classifying, extracting. A model that costs a twentieth of the frontier price handles them at 95% of the quality (LMSYS, 2024). The trick is deciding per task, not per person, and doing it automatically.

Open-source models. Run on your own cloud or ours, they cost a fraction of a lab's API price for the same class of task and keep the data inside your boundary. The right answer for a large share of routine work, and the wrong answer for some. A gateway makes the choice per request.

Seats and duplicates. 36% of software licences go unused (Zylo, 2026). Two teams paying for the same capability is the norm, not the exception.

£16,000

a month back for a 300-person firm, at standard routing. £192,000 a year.

GenWA calculator · assumptions above

£19,900

a month with open-source models in the mix. £238,000 a year, at the 72% cap.

GenWA calculator · assumptions above

PAY FOR OUTCOMES, NOT TOKENS

Token consumption benefits the labs. Someone independent of them should be optimising your spend, and should be paid on what it saves you. Run your own numbers at genwa.co.uk/#/businesscase.

Rules as controls, through one gateway

A policy that lives on the intranet is a document. A policy that sits in the path of every request is a control. The difference is a gateway: one door between everything that uses AI and every lab that provides it.



WHAT THE GATEWAY DOES ON EVERY REQUEST

Identifies who or what is asking, on which account, for which purpose. **Checks** the data against your rules: personal data, client material, special-category, commercial IP. **Chooses** the model that satisfies the task and the policy at the lowest cost. **Records** the request, the decision and the outcome, so page 19 writes itself. **Stops** what should not happen, and says why.

Because the rules are written about tasks, data and risk rather than about a vendor, a switch from one lab to another is a test and a change record, not a rewrite. That is what detaches your policy from the turbulence on page 14.

FOUR RULES, AS CONTROLS

THE RULE	THE CONTROL
No customer personal data in consumer AI tools	Gateway detects personal data and blocks or redacts before the request leaves
Premium models only where the task needs them	Router sends routine work to the standard tier; premium requires a task class
Decisions about individuals need a human	Outputs tagged “decision” are held for review; the reviewer’s name is logged
Agents act within a budget and a scope	Per-agent spend cap, scoped credentials, kill-switch, full action log (NCSC, 2026)

The one-page board report

Risk and return side by side, in the units the board already uses. Two-thirds of boards have limited or no AI knowledge and at a third AI is not on the agenda at all (Deloitte, 2025). Boards that discuss AI at every meeting are far more likely to sit in high-return organisations: 63% against 13% (Protiviti, 2026).

1 · EXPOSURE

Number of AI tools in the register; new this month; retired this month.

Company accounts versus personal accounts, as a count and a trend.

Data incidents and policy exceptions: how many, what kind, how resolved.

Any tool or use touching personal, client or special-category data without a documented basis.

2 · SPEND

Total AI spend this month, to the pound, against budget and last month.

Spend by team and by tool; the top five lines. Forecast for next quarter and the assumptions behind it.

Premium-model share of cost, and what it was last month.

Unused seats and duplicate subscriptions, with the saving taken.

3 · MODELS

The approved model list and any changes this month: additions, retirements, price or terms changes.

Which workflows were re-tested after a model change, and the result.

4 · COMPLIANCE AND RETURN

Policy status: published, trained, enforced where. Training coverage as a percentage of staff.

Named owner for each AI system that makes or shapes decisions about people.

The three use cases under measurement: the number, the trend, and what has been stopped as a result.

THE EVIDENCE PACK

Behind the page sits the record: every request, decision and outcome from the gateway, mapped to UK GDPR, the DUAA safeguards, the ICO code and ISO/IEC 42001 clauses. When the auditor, the insurer or a customer asks, it is generated, not assembled.

63% vs 13%

high-return organisations among boards that discuss AI at every meeting, against those that rarely do.

Protiviti, 2026

MOVE 05 · SCALE · TEMPLATE

The 90-day plan

Three phases. See first, because everything else is guesswork without it. Then cut and contain. Then prove, and scale on what holds.

WEEKS 1-3 · SEE

WEEK	ACTION	OWNER	DONE
Week 1	Pull every AI line from three months of expenses, card statements, subscriptions and cloud bills. Name the team behind each. Send a short, non-punitive survey: which AI tools do you use for work, on which account?		☐
Week 2	Start the register (page 16). Add AI features inside existing software. Map what data goes into each tool. Check the default model in every paid tool and in any internal code.		☐
Week 3	Write the findings on one page: tools, accounts, spend, data exposure, duplicates, and the top three risks. Take it to the executive team. Agree who owns AI from here on.		☐

WEEKS 4-8 · CUT AND CONTAIN

WEEK	ACTION	OWNER	DONE
Week 4	Choose the approved tools. Give everyone a company-owned account on at least one. Announce it, with what it may and may not be used for.		☐
Week 5	Write the AI policy in plain English, two pages at most: allowed uses, banned data, who approves what, what is logged. Align it to UK GDPR and the DUAA; note where ISO/IEC 42001 will ask for more.		☐
Week 6	Set the default model per task class. Make premium a deliberate choice. Decide where AI must not be used at all.		☐
Week 7	Move work from personal to company accounts. Retire or merge the duplicates. Cancel unused seats. Take the saving.		☐
Week 8	Put the policy into the path of use: a gateway, access controls, data rules, approval steps for customer-facing output.		☐

WEEKS 9-12 · PROVE AND SCALE

WEEK	ACTION	OWNER	DONE
Week 9	Set a monthly AI budget by team and alerts for unusual usage. Agree one number a month for total spend, and who produces it.		☐
Week 10	Pick three AI uses and define how you will measure their return. Set a review date and a stop criterion for each.		☐
Week 11	Create a shared, attributed library of working prompts, workflows and agents. Deploy the first agent on the controlled foundation, with a budget and a scope.		☐
Week 12	Produce the first board report (page 19). Re-run the self-check (page 21) and compare. Set the date for the next quarter's model review.		☐

PRINTABLE

The seven-question self-check

Be strict. “Yes, with evidence” means you could show the document, the report or the screen today, not that you believe it to be true.

1	Can you list every AI tool in use in the business, including tools on personal accounts and AI features inside software you already own? SEE PAGES 8, 10, 16	☐ Yes, with evidence ☐ Partly ☐ No
2	Do you know last month's total AI spend, to the pound, across expense claims, subscriptions, cloud bills and API usage, and which team incurred each part? SEE PAGES 8, 16	☐ Yes, with evidence ☐ Partly ☐ No
3	Does anyone in the business decide which AI model is used for which kind of task, or is it left to each person and each tool's default? SEE PAGES 9, 17	☐ Yes, with evidence ☐ Partly ☐ No
4	Does every employee do work-related AI on a company-owned account that you could see, manage and switch off? SEE PAGES 10, 18	☐ Yes, with evidence ☐ Partly ☐ No
5	Is there a written AI policy, and is it enforced somewhere other than the intranet: at the point of use, with a record of exceptions? SEE PAGES 14, 18	☐ Yes, with evidence ☐ Partly ☐ No
6	Could you show a regulator, an insurer or a customer what data goes into your AI, what comes out, and who approved it? SEE PAGES 14, 19	☐ Yes, with evidence ☐ Partly ☐ No
7	Do you know which AI uses have produced a measured return, which have not, and what you have stopped as a result? SEE PAGES 13, 19	☐ Yes, with evidence ☐ Partly ☐ No

SCORING · YES, WITH EVIDENCE 2 · PARTLY 1 · NO 0 · MAXIMUM 14

12–14

In control. Your job now is to keep it that way as the models and the rules change. Page 14 and the board report on page 19 are where to focus.

7–11

Gaps. You can probably answer for the AI you bought and not for the AI that walked in. Start with the register on page 16 and the first three weeks of the plan.

0–6

Exposed. You are paying for AI you cannot see and carrying risk you cannot evidence. Do the “this week” actions on pages 8 to 13 before anything else.

The score is not a judgement; it is a starting line. The same seven questions, scored automatically with a written report, are the ten-minute online health check at genwa.co.uk/#/healthcheck.

TAKE-AWAYS · TEAR OUT

Seven things to take into the next board meeting

1

Ask for the number.

Last month's total AI spend, to the pound, across expenses, subscriptions, cloud and API. If nobody can produce it in a week, that is the finding.

2

Own the accounts.

Every employee on a company-owned AI account you can see and switch off. Personal accounts are exposure you cannot measure.

3

Make premium a decision.

Set the standard tier as the default in every tool and every line of code. Premium by choice, with a reason.

4

Detach policy from vendors.

Write rules about tasks, data and risk, not about a model. Then a model change is a test, not a rewrite.

5

Put the rules in the path.

A policy on the intranet is a document. A gateway that checks every request is a control. Only one of them will satisfy the ICO.

6

Generate the evidence.

If the regulator, insurer or a customer asked tomorrow, what would you show them? The answer should be a report, not a project.

7

Scale only what holds.

Agents multiply whatever foundation they sit on. Deploy them with a budget, a scope and a kill-switch, on a register that already exists.

THE NUMBERS TO REMEMBER

71%

of UK staff have used AI their employer never approved.

56%

of AI cost goes to premium models.

26%

of companies fully track what AI costs them.

4×

the rate of established return where AI cost is fully visible.

Sources on page 24. Score yourself on page 21, or in ten minutes at genwa.co.uk/#/healthcheck.

HOW GENWA HELPS

All the upside of AI. None of the surprises.

GenWA is the control layer between your business and every AI lab: one gateway, your rules, any model. It gives the board back what it lost when AI walked in: visibility, control and evidence.

Everything in this playbook can be done by hand. Most of it should be done by software, because it has to be done every day, and the models change every few months. GenWA's six product families map onto the five moves and the advisory work around them.

FAMILY	WHAT IT GIVES YOU
See	Discover, Watch, Cost. The register built from your data in 48 hours and kept live.
Cut	Routing to the cheapest adequate model, open source included; seats and duplicates retired.
Contain	Gateway, Rules, Change Control. Your policy enforced on every request; a model switch as a governed event.
Prove	Ledger, Evidence Packs, Compliance Mapping. UK GDPR, DUAA, ICO code and ISO 42001, generated from the record.
Scale	Agents and agent teams, trained on your sector's rules and roles, on your cloud or ours, within a budget.
Advance	Audit, Policy and Controls, Training, Managed Operations.

START WITH THE AI HEALTH CHECK

Ten minutes online. The seven questions, scored, with a written report.

48 hours with your usage data. Upload an export; usage only, never content, anonymised in your browser. You get the register, the savings and the exposures.

Half a day on site. A consultant, your executive team, the findings on one page.

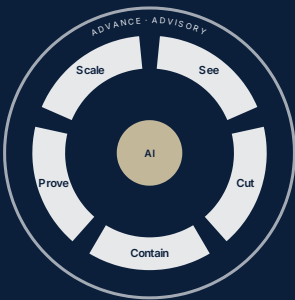
48h

from your usage data to a register, a savings figure and an exposure list.

£0

for the health check. No obligation.

genwa.co.uk · hello@genwa.co.uk



Sources

Official statistics and regulators first; vendor-commissioned surveys and vendor telemetry are marked as such. Where two sources disagree we have used the more conservative figure.

1. ONS, Artificial intelligence in UK businesses, July 2026.
2. British Chambers of Commerce, AI adoption survey, March 2026.
3. DSIT, AI Adoption Research, February 2026.
4. DSIT, Cyber Security Breaches Survey 2025/26, April 2026.
5. Bank of England and FCA, Artificial intelligence in UK financial services, 2024.
6. CIPD, Labour Market Outlook, Autumn 2025; CIPD, January 2026.
7. Microsoft UK / Censuswide, Work Trend Index UK, 2025 (vendor survey).
8. KPMG UK, AI cost visibility survey, July 2026 (n=116).
9. KPMG, Global AI Pulse, 2026.
10. Ramp, AI Index, April 2026 (vendor data).
11. Epoch AI, LLM inference price trends, 2025.
12. Mavvrik, AI cost survey, 2026 (vendor survey).
13. Zylo, SaaS Management Index, 2026 (vendor data).
14. Bloomberg, Uber AI budget, 2026.
15. Gartner, AI cost estimation, 2024; Gartner, agentic AI project cancellations, June 2025.
16. a16z, How 100 enterprise CIOs are building and buying gen AI, 2025.
17. Anthropic, Claude Code documentation on cost, 2026.
18. OpenAI, GPT-5 routing usage note, August 2025.
19. LMSYS, RouteLLM, 2024.
20. Cyberhaven, AI Adoption and Risk Report, 2025 and 2026 (vendor telemetry).
21. MIT NANDA, The GenAI Divide, 2025.
22. IBM, Cost of a Data Breach Report UK, 2025 and 2026 (vendor research).
23. Netskope, Cloud and Threat Report, 2026 (vendor data).
1. Salesforce / MuleSoft, Connectivity Benchmark, 2026 (vendor survey).
2. BetterUp and Stanford Social Media Lab, in Harvard Business Review, 2025.
3. Upwork Research Institute, 2024 (vendor survey).
4. McKinsey, The State of AI, 2026.
5. Microsoft, Work Trend Index, 2026 (vendor survey).
6. KPMG and University of Melbourne, Trust in AI, 2025.
7. Cornerstone OnDemand, UK workforce survey, 2025 (vendor survey).
8. IBM Institute for Business Value, CEO Study, 2025.
9. ABBYY, UK IT leaders survey, 2024 (vendor survey).
10. WRITER, Enterprise AI adoption survey, 2026 (vendor survey).
11. PwC, Global CEO Survey, January 2026.
12. S&P Global Market Intelligence, 2025.
13. Avalara, CFO survey, 2026 (vendor survey).
14. EY, AI cost pulse, 2026.
15. DLA Piper, Data (Use and Access) Act commencement note, 2026.
16. ICO, AI and automated decision-making code of practice, 2026; SI 2026/425.
17. Trowers and Hamlin, ICO enforcement review, 2026.
18. UKAS, ISO/IEC 42001 accreditation, January 2026.
19. NCSC, Guidance on AI agents, 2026.
20. Deloitte, Board practices and AI, 2025.
21. Protiviti, Board perspectives on AI, 2026.
22. Dario Amodei, We Must Pace the Frontier, 14 September 2026.
23. Microsoft and LinkedIn, Work Trend Index, 2024 (vendor survey).

© GenWA Ltd 2026. This playbook is general guidance and not legal advice. Figures are correct to the best of our knowledge at the date of publication and are attributed to their sources; readers should check current guidance from the ICO, the FCA and the NCSC before relying on regulatory statements. Second edition, September 2026. [confirm: company registration number and registered office]



All the upside of AI. None of the surprises.

[GENWA.CO.UK](https://genwa.co.uk)