A SHORT BRIEF FOR FOUNDERS AND AGENCY OWNERS

The Shadow AI Gap

Your team has already adopted AI. You did not approve it, you cannot see it, and it may be outperforming the programme you did approve.

THE FINDING NOBODY EXPECTED

THE OFFICIAL PROGRAMME

Procured. Governed.
Rolled out with a plan.

Bespoke enterprise systems

Budget approved and tracked

Orders of magnitude more expensive

THE SHADOW ONE

Unapproved. Unseen.
Paid for personally.

A general tool at about $20 a month

Chosen by the person doing the work

Often the better return of the two

MIT found personal AI often delivers better ROI
than the formal initiative sitting beside it.

So the instinct to ban it is the wrong instinct.

Your team is running a live experiment and quietly finding out what actually works.

The productivity is real. The governance is missing. Only one of those is a problem.

MIT Project NANDA, The GenAI Divide: State of AI in Business, 2025

Paul Prado Pacardo

Senior Executive Assistant · Operations and Project Manager

me.kabuhayan.app

linkedin.com/in/paulpacardo

pacardopaul18@gmail.com

THE GAP BETWEEN THE LINE ITEM AND THE DESK

Forty percent authorised it. Over ninety percent are doing it.

This is not a story about rule-breaking. It is a story about a workforce that found something useful and did not wait for procurement to catch up.

40%

of companies report purchasing an official large language model subscription

MIT Project NANDA, 2025

90%+

of surveyed companies have staff regularly using personal AI tools for work

MIT Project NANDA, 2025

78%

of AI users bring their own AI to work. At small businesses it rises to 80 percent

Microsoft and LinkedIn, 2024

WORTH NOTING

Small companies are more exposed, not less

Microsoft's Work Trend Index found bring-your-own-AI running at 80 percent in small and medium businesses against 78 percent overall. The instinct that you are too small for this problem is precisely backwards.

AND

Around half will not tell you they used it for anything important

Same study, 31,000 workers across 31 countries. People conceal it because they expect to be told to stop, not because they think they are doing harm.

METHOD NOTE

MIT NANDA is preliminary findings from a research group, not a peer-reviewed paper

300 public deployments reviewed, 52 executive interviews, 153 survey responses. Credible and widely cited, and I am telling you its status rather than letting you assume.

WHAT THIS ACTUALLY MEANS

You do not get to decide whether your business uses AI. That decision was made, individually, by people trying to finish their work. You only get to decide whether you can see it.

WHAT IT COSTS WHEN NOBODY IS LOOKING

The same breach costs $670,000 more.

From the most rigorous dataset available on this: IBM's annual breach study, conducted with the Ponemon Institute, across 600 breached organisations in 16 countries.

+$670k

average cost premium where shadow AI was involved: $4.63M against $3.96M

IBM with Ponemon, 2025

97%

of affected organisations had no AI access controls in place at all

IBM with Ponemon, 2025

65%

involved customer personal data, against 53 percent globally

IBM with Ponemon, 2025

READ IT RIGHT

That is a premium, not your bill

The $4.63M figure comes from organisations far larger than yours. The useful part is the gap between the two numbers, which says ungoverned AI makes an incident materially worse, not that a small firm faces millions.

THE REAL EXPOSURE

For a service business it is client confidentiality, not a headline breach

The realistic failure is a contract, a client list or an unreleased campaign pasted into a consumer tool with training enabled. No hacker required, no alarm sounded, and nothing you can retract.

NINETY-SEVEN

The controls figure is the actionable one

Almost nobody had any AI access controls. That is not a sophisticated attack surface. It is an open door, and closing it is administrative rather than technical.

THE POINT OF THIS PAGE

This is not an argument for banning it. It is an argument for knowing what is already happening, which costs you nothing and is the only thing a ban would have achieved anyway.

THE FINDING THAT SHOULD CHANGE YOUR RESPONSE

The unapproved tool is often the one that works.

Everything written about shadow AI treats it as a risk to be eliminated. MIT's research suggests it is also the most honest signal you have about what actually helps.

The official programme

Bespoke enterprise systems, procured properly, budgeted and governed, rolled out with a plan. MIT's wider finding was that around 95 percent of these pilots produced no measurable impact on profit and loss.

The shadow one

A general-purpose tool at roughly twenty dollars a month, chosen by the person doing the work, paid for personally. MIT found these often deliver better return than the formal initiative running alongside them.

WHY

The person who chose it is the person doing the task

No procurement committee, no requirements document, no vendor demo. They tried it on real work on Tuesday and kept it because it helped. That is a shorter feedback loop than any pilot you could run.

WHICH MEANS

Your staff have already run the evaluation you were planning to commission

Unfunded, unmanaged, and with better information than a vendor pitch. The results are sitting in people's browser histories and nobody has ever asked to see them.

THE MISTAKE

Banning it deletes the evidence and keeps the risk

The usage does not stop. It moves to personal devices, where you have less visibility, not more. You lose the signal and retain the exposure.

THE REFRAME

The productivity is real. The governance is missing. Only one of those is a problem, and it is not the one most companies act on first.

A NOTE ON THE NUMBERS IN THIS FIELD

Most statistics here are sold by whoever fixes it.

Security vendors measure how much sensitive data goes into AI tools, and also sell the software that stops it. That does not make them wrong. It makes them worth naming, and it makes some figures worth leaving out.

What I left out

The widely quoted percentages for how much pasted data is sensitive. One vendor's own reporting moves from about 11 percent to 35 percent to nearly 40 percent across successive years, measuring somewhat different things each time. A figure that unstable is not a measurement, it is a trend line with a marketing department.

What I kept

MIT Project NANDA for prevalence and the ROI inversion. IBM with Ponemon for breach economics, because it interviews real breached organisations and publishes its method. Microsoft and LinkedIn for bring-your-own-AI, because the sample is 31,000 people and the fielding agency is named.

ALSO FLAGGED

Microsoft sells AI

Their Work Trend Index is large, well documented and useful, and they have an obvious commercial interest in showing enthusiastic adoption. Both things are true at once and you should hold them together.

ALSO FLAGGED

MIT NANDA is preliminary, not peer reviewed

It has been publicly criticised since publication. I use it because its findings on this specific point match what I have watched happen inside real businesses, and I would rather say that than pretend it is settled science.

THE STANDARD I AM APPLYING

Name the study, the date and the sample. State the interest of whoever published it. A number without a method is an advertisement, and you should treat this brief the same way.

WHAT TO ACTUALLY DO ABOUT IT

An approved list beats a ban, every single time.

You need three things, none of which require software, a consultant, or a policy document nobody will read.

STEP 01

Ask, without consequence

Tell the team plainly that nobody is in trouble, then ask what they are using and what it actually helps with. You will get honest answers exactly once, and only if the first sentence is credible. This is the whole audit.

STEP 02

Approve two, pay for them

Pick the two tools that came up most, buy the business tier, and put everyone on them. The paid tiers keep your data out of training by default, which is most of the risk removed for the price of a couple of subscriptions.

STEP 03

Write the rule on one line

Not a policy document. One sentence people can recall under pressure, plus a named person to ask when something is borderline. If it takes a paragraph, it will not be followed.

THE RULE, AND IT FITS ON A CARD

Never paste anything you would not email to the wrong client by accident. **If you are unsure, ask before you paste, not after.**

WHY THIS ORDER

Asking first is what makes the rest work. Approve tools before you understand what people need and you will buy the wrong two, and the shadow usage carries on beside them.

EVIDENCE, NOT ADJECTIVES

I use these tools daily, and I hold a line about it.

I am not writing this as an observer. AI is part of how I work, which is also why I am specific about where it does not belong.

Where I do not use it

Anything a stakeholder should hear from a person. Final judgment on priority or escalation. Confidential material outside approved tools. Numbers I have not reconciled myself. Apologies, bad news and difficult conversations.

Where it earns its place

First drafts I then rewrite in the executive's voice. Compressing long threads into a decision list. Turning transcripts into owners and dates. Comparative research before a vendor decision. Writing and debugging the automation scripts that do the rest.

CERTIFIED	**Google AI Essentials, and Anthropic AI Fluency including Model Context Protocol** Framework and Foundations, AI Capabilities and Limitations, Model Context Protocol Advanced Topics, and AI Automation with Claude. Learning the failure modes deliberately, not just the prompts.
APPLIED	**40+ automation rules, 9 scripted jobs, 8 integrations built with AI assistance** Across Jira, ScriptRunner and Google Workspace, in a live agency. Used to analyse workflows, write and debug scripts, and pressure-test logic before deployment.
THE PRINCIPLE	**Automate a process only after it is documented, owned and measured** Pointing any tool at an undefined workflow produces the same mess faster. That judgment matters more than the tooling, and it is the thing an approved list cannot give you.

THE LINE I HOLD

AI drafts. I decide. Anything that reaches a human has been read by one first.

HOW TO START

Ask the question, and mean the amnesty.

The whole thing turns on one conversation, and on whether people believe the first sentence of it. Everything after that is administration.

OPTION 01

Run the amnesty yourself

One message: nobody is in trouble, tell me what you are using and what it helps with. Expect to be surprised by at least one answer, and expect at least one person to have solved something you were about to pay for.

OPTION 02

Let me write the rule and the list

The approved tools, the one-line rule, the named person to ask, and the paid tiers configured so your data stays out of training. A week, and it does not need a policy document.

OPTION 03

Start with the highest-risk workflow

Whichever process touches client confidential material most often. Document what may and may not go into a tool for that one workflow, then extend the pattern outward.

WHAT I WOULD EXPECT YOU TO FIND

More usage than you thought. At least one tool genuinely worth paying for. And one person quietly doing something clever that nobody else in the business knows about.

Available now for remote work

Chief of Staff · Operations · Senior Executive Assistant · Project Manager

Tell me what came back from the amnesty and I will tell you what I would approve first.

pacardopaul18@gmail.com
linkedin.com/in/paulpacardo
me.kabuhayan.app

Sources: MIT Project NANDA, The GenAI Divide: State of AI in Business 2025; 300 public deployments reviewed, 52 executive interviews, 153 survey responses; preliminary findings, not peer reviewed. IBM Cost of a Data Breach Report 2025, conducted with Ponemon Institute; 3,470 interviews across 600 breached organisations in 16 countries. Microsoft and LinkedIn Work Trend Index 2024; 31,000 workers across 31 countries, fielded by Edelman Data and Intelligence. Deliberately excluded: vendor figures for the proportion of pasted data that is sensitive, which vary from roughly 11 to 40 percent across successive reports by the same publisher. Figures are third-party research, not results I am claiming.