GOVERNANCE-FIRST AI · REGULATED OPERATIONS

Automating the closed-system last mile.

A governance-first approach to AI in regulated operations: why the model was never the hard part, and what it takes to safely reach the work no model can touch.

THE THESIS

The model was never the hard part. The cost sits in the last mile no model can reach.

For two years the conversation about AI in operations has been a conversation about models. Which model is most capable. Which is cheapest per token. Which will be obsolete by the next release. It is the wrong conversation. In a regulated operation the model is the commodity. The hard part, the part that decides whether anything works in production, is everywhere the model isn't.

Picture the real shape of the work. A claim arrives by email with a PDF attached. To progress it, someone logs in to a portal protected by multi-factor authentication (MFA), reads three screens, copies four fields into a legacy system that has no application programming interface (API), updates a shared spreadsheet, and emails a broker. None of that is a reasoning problem. All of it is a last-mile problem: closed systems, human credentials, no clean integration point, and a regulator who expects to see who did what and why.

This is the closed-system last mile. It is where the hours actually go, and it is precisely the work that most AI offerings route around, because it is hard, regulated, and does not demo well.

> Everyone sells the part that demos well. The cost lives in the part that doesn't: the portal, the legacy screen, the inbox, the spreadsheet.

The gap matters because operations leaders are being sold a finished product and handed an unfinished one. A capable model plus a clean API is a weekend project. A capable model that can operate an MFA-gated portal under supervision, leave an audit trail a regulator will accept, and stop before it crosses a line that must never be crossed, is a different discipline entirely.

That discipline is what this paper is about. We go to the last mile on purpose. The rest of these pages set out what you are really choosing between, the method we use to get there safely, and the evidence that it returns real time without taking on unacceptable risk.

CLOSED	MANUAL	REGULATED
MFA-gated portals and legacy systems with no API.	Inboxes and spreadsheets that carry the real workload.	Every action must be explainable, auditable and reversible.

WHAT YOU'RE REALLY CHOOSING BETWEEN

Five honest options. Each one stops somewhere short of the last mile.

No option is wrong for every problem. The question is where each one stops, and whether the work it cannot reach is the work that is actually costing you.

01

Horizontal tools

Off-the-shelf copilots and workflow apps. Fast to switch on, genuinely useful for generic tasks. They stop at your perimeter: they cannot log in to your MFA-gated portal or drive a legacy system with no API, so the last mile stays manual.

02

Big consultancies

Strategy, slideware and a large team. Strong on transformation narrative. They tend to stop at the recommendation, priced by the hour, with the integration risk and the closed-system reality handed back to you to resolve.

03

Generic AI consultants

Comfortable building a demo on a clean dataset. They stop where the regulation starts: governance, per-action approval, audit trails and the red lines that must hold are treated as someone else's problem, not part of the build.

04

Hiring in-house

Full control and durable knowledge, if you can hire it. It stops at capacity and time: applied-AI, regulatory and operations skill in one team is scarce and slow to assemble, and the closed-system work is unforgiving to learn on the job.

05

Doing nothing

A legitimate choice, and sometimes the right one. It stops being safe when the manual last mile is quietly costing real money. One client was spending roughly 900 operational minutes a day on it, about £120,000 a year.

WHERE NORTH STACK FITS

We take the work the other four route around and the fifth pays for: safely automating the closed-system last mile, with a person approving every critical action and a full audit trail behind it. We also say no in writing when a workflow shouldn't be automated.

> The test isn't which option is best. It's whether the work an option can't reach is the work that's costing you.

THE METHOD

Prove the hard part first. Climb one rung at a time. Keep a person in command.

Audit-first, before anything is built

Every engagement opens with a fixed-fee operations review. We map the operation, score each workflow, and hand back a costed, sequenced plan, including the work we would tell you not to automate. No build is priced before the hard part is proven on your real data, not a sanitised sample. If a workflow shouldn't be automated, we say so in writing.

The North Stack Ladder

We don't leap to autonomy. We climb five rungs deliberately, stopping at the point where the return is real and the risk is controlled. Most of the value arrives well before the top.

01 **Mapped** the operation is documented and scored.

02 **Assisted** AI drafts; a person decides.

03 **Automated** routine steps run with approval.

04 **Integrated** workflows connect across systems.

05 **Orchestrated** supervised agents run end to end.

Supervised agents that operate closed systems

When we reach the last mile, agents work the closed systems the way a trained operator would: signing in to the MFA-gated portal, reading the legacy screen, drafting the email, updating the record. They never act alone on anything that matters. A person approves every critical action, and every step is written to a complete, exportable audit trail.

GOVERNANCE, BUILT IN, NOT BOLTED ON

- Human-in-the-loop on every high-stakes step.
- Per-action approval for agents.
- Field-level confidence scores on extracted data.
- Complete, exportable audit trail.
- Role-based access control.
- A failure and incident playbook.

RED LINE

We never automate payments.

RED LINE

We never automate bank-detail changes.

RED LINE

A hard human gate sits before any regulator submission.

PROOF ON REAL OPERATIONS

The method returns real time, without taking on unacceptable risk.

~60%	~70%	~900	~3 days
less manual effort, with quotes more than 50% faster, live in about six weeks — PMD Finance	less document-review time and reporting about 45% faster, on their own insolvency CRM — Marshall Peters	operational minutes a day returned to one client, about £120,000 a year	faster board actions — Norcis. Further work with Ziani and Hippo

How to start

Start with the fixed-fee operations review. There is no obligation to build, and no build is priced until the hard part is proven on your data. You leave with a clear map of the operation, a workflow-by-workflow score, and a costed, sequenced plan, including an honest list of what we would not automate. If the answer is to do less, we will tell you, in writing.

It is the cheapest way to find out whether the closed-system last mile is costing you what we usually find it costs, and the safest way to do something about it.

The team behind it

A UK studio of five, remote-first, built deliberately for regulated operations:

- Lewis Leach, former UK Parliament counsel.
- Matthew Place, formerly Lloyds and HSBC.
- Myles Langstone, formerly the ABI.
- Cho Yin Yong, applied AI, two patents, University of Toronto.
- Dr Kasra Milani, MD.

> We say no in writing. No build is priced before the hard part is proven on your real data.