

Not all bad data is the same kind of bad.

The readiness checklist flattens four different problems into one word and prices them the same. One a reasoning model absorbs. One it catches against a working picture of the world. One **nothing** catches — and its fix is upstream, not downstream. And one is not a data problem at all — it is the record of a decision nobody made. **Sort the mess before you buy anything to clean it.**

THE KIND OF WRONGWHAT ACTUALLY HANDLES IT — AND WHAT DOES NOTVERDICT

Interpretive mess
Inconsistent formats, renamed fields, unstructured documents, the exception no rule covered

The model reads around it — RPA snapped when a column got renamed. **Lever: model selection** — match capability to how readable one item is without a human, not to how many there are.

ABSORBED

Implausible error
An invoice for consulting services from a vendor whose registered activity is scaffolding

Caught by reading internal consistency and expected distributions against a working picture of the world. A rules engine catches only what somebody configured — a difference in *coverage*, not cleverness. **Lever: the halt boundary and its sensitivity dial**, which someone owns, because every turn of it adds queue volume.

CAUGHT

Plausible error
The stale price that is still a normal price; the valid address for a customer who moved

Nothing catches this — and a human reviewer misses it too, because they see a normal price and a valid address. Internally consistent, distributionally ordinary, and wrong. **No detection lever exists.** The only lever is upstream, in the process that creates the record: confirm-your-details at login every six months, a flag on any price untouched for twelve. **Freshness as a control, not accuracy as a target.**

PROCESS REDESIGN REQUIRED

Unmade decisions
Five spreadsheets, five revenue figures, and no record of anyone deciding which one is the number

Not a data problem — the answer is not in the data at all. A bounded agent halts and escalates. An unbounded one fills in: confident, plausible, silent. **Lever: a named owner** with a scoped veto, a funded budget line and an unfiltered route upward.

A HUMAN DECIDES

THE NEW FAILURE MODE IS QUIETER THAN THE ONE IT REPLACED

RPA broke loudly. **An agent guessing at an unmade decision looks exactly like an agent doing its job.** So the boundary has to be enforced where the model cannot reach it — a tool it cannot call, a write it cannot make, a reconciliation check that hard-fails before the agent is asked what it thinks. A boundary written into a prompt is not a boundary; it is one more input the model weighs.

THREE THINGS TO SETTLE BEFORE IT RUNS

- 01

Which of the four is actually in front of you? Three can be handled where the agent runs. One can only be handled upstream, in the process that creates the record.
- 02

Where is the boundary enforced — in code, or in a prompt? Specify it as an allowlist of what the agent may settle alone, with everything outside it halting by default.
- 03

Who answers the queue, and by when? Name them, fund the line, and put a clock on the route up: unanswered inside five working days, it goes to the risk committee unfiltered.

Reasoning resolves ambiguity.
It does not resolve unmade decisions.

Source: Fredrik Lindstrom, *AI Doesn't Need Perfect Data*, 21 August 2026 — sections “What an agent absorbs, and what it cannot” and “Build the agent to escalate”. The four-way split is this sheet’s framing of that argument and is not a taxonomy drawn from a cited study. **Structural over prompt-layer:** the preference for rule-based controls enforced outside the model is stated in Singapore IMDA’s *Model AI Governance Framework for Agentic AI*. **Human oversight:** EU AI Act Article 14 requires that oversight can be exercised, and Article 26(2) puts the competence, training, authority and support to exercise it on the deployer — a four-second review satisfies neither. **The upstream lever** on the plausible-error row — periodic re-confirmation at login, staleness flags on unmaintained fields — is this sheet’s addition and does not appear in the source article. None of the levers above cleans a data estate; three route each kind of wrong to the thing that can act on it, and the fourth prevents the record rather than catching it. Row groupings and verdict labels above are this sheet’s reading of the cited source, not the source’s own framing.

FOLLOW FREDRIK LINDSTROM FOR MORE

Fredrik Lindstrom

AI and cybersecurity governance for boards, CISOs and senior leaders. 25+ years in enterprise security.

A practitioner to guide you.