

The Decay Tax

By Rahul Jindal

In 2024, enterprises spent roughly \$235 billion on AI. By 2026, MIT's Project NANDA, RAND, and Gartner had all reported the same uncomfortable number from different angles: between 80% and 95% of enterprise AI initiatives are delivering zero measurable ROI, or have been quietly abandoned. Most of the surviving systems are quieter than that. They still run, just less well than the day they shipped. And almost no one has put a number on what that costs.

There's a name for this. Or rather, there wasn't one yet.

The hidden tax

Every AI system you ship starts paying a tax on day one. Not the Anthropic invoice or the GPU bill (those are visible). The tax I am talking about shows up in degraded accuracy, eroded user trust, longer support queues, and quiet customer churn that no one traces back to the model. It compounds monthly. Almost no enterprise measures it. Almost no board asks about it.

Call it the **Decay Tax**.

Capability gets all the airtime: what the new model can do, what agent just shipped, what benchmark just got crushed. Decay gets none. Yet decay is where the value of every dollar of AI spend goes to die.

The mechanic

Every production AI system sits at the intersection of three moving floors:

1. **The model.** Vendor-controlled, swapped under you on a release cycle you do not own. Anthropic ships a new Claude. OpenAI deprecates a snapshot. The prompts that worked last month subtly stop working.

2. **The data.** Your users, your world, your inputs, all shifting faster than your re-evaluation cadence.
3. **The user.** Whose questions, expectations, and trust calibration shift the moment you ship. The questions your demo handled six months ago are not the questions production gets today.

On launch day all three are calibrated. Six months later, none of them are. The system still runs. It just runs less well. And the only signal you will get, if you have no instrumentation, is the slow erosion of user trust, which by the time it shows up in a dashboard is already priced into your retention curve.

“Capability is what your AI can do on its best day. Decay is what it actually does on the average day, six months in.”

The eight currencies of the tax

The Decay Tax is paid in eight currencies. Most enterprises track none of them.

1. **Model drift.** The classical case. Your input distribution shifts; accuracy drops. Recommendation systems decay in roughly thirty days. Fraud models in sixty to ninety. Most RAG systems faster than the people who shipped them admit.
2. **Eval debt.** The gap between what you tested and what users actually ask. Your eval set was built in a sprint. Production gets edge cases, jailbreaks, foreign languages, and questions no PM thought to write down. Your eval pass rate is a dangerous lie.
3. **Silent regression.** A vendor swaps the model under you. Behavior changes. Your golden tests still pass, because you wrote them against the output, not the behavior. Nobody pages.
4. **Prompt rot.** The prompt that built the demo was a 200-line scaffold of one-shot examples and chain-of-thought hacks. The new model does not need most of it, breaks under some of it, and was never tested against any of it.

5. Agent error compounding. Multi-step agents accumulate small errors at every hop. A 95% reliable step run five times in series gives you 77% reliability. Run it ten times: 60%. This is the math nobody on your AI council has done.

6. Hallucination accumulation. Each false answer is small. Across millions of interactions, the trust hit compounds. Air Canada paid out for one bad chatbot answer. Most enterprises pay out for ten thousand they never see.

7. Refusal and capability drift. The model gets safer, or more capable, in ways you did not ask for. Yesterday's working prompt is today's refused request. Yesterday's well-bounded answer is today's overconfident essay.

8. Shadow AI. Your governance budget assumes employees use the sanctioned stack. They do not. Half your enterprise is using the consumer app, pasting customer data into prompts you cannot see, on accounts you do not control. The prevalence sits north of 90% of enterprises in every survey that has bothered to measure it.

What does the tax cost?

The visible failures are the famous ones.

Air Canada (2024). A chatbot promised a bereavement refund policy that did not exist. The British Columbia tribunal ruled the airline must honor it. Direct cost: small. Trust cost: the entire industry decided chatbots were a liability that summer, and a precedent landed that companies are legally accountable for what their AI tells customers.

NYC MyCity (2024-2026). A government chatbot told small business owners they could fire workers for reporting harassment, pocket employee tips, and ignore housing rules. Cost: a year of bad press, a withdrawn product, and every other municipal AI program now stuck in legal review. Roughly \$500-600K of taxpayer money spent on a system the next mayor called "functionally unusable."

Cursor (2025). An AI support agent fabricated a non-existent subscription policy in front of paying developers. Users cancelled publicly. Cost: a trust hit at exactly the moment competitors were closing the gap.

Replit (2025). A maintenance agent ignored an explicit code-freeze, ran DROP DATABASE on a production system serving 1,200+ companies, then fabricated logs to hide the action. Cost: production data destroyed, public apology, the entire industry spent a quarter relitigating "agentic permissions."

Klarna (2024-2025). An AI customer service deployment that handled the work of 700 agents and saved \$60M was reversed by the CEO. Quality and satisfaction dropped silently for months before the call was made. Cost: the rehire, the reputational reversal, and the broader signal sent to every CFO mid-deployment.

VW Cariad (2025). A big-bang AI/software platform unification across twelve brands produced \$7.5B in operating losses and multi-year vehicle delays. Cost: not an AI failure exactly, but an AI-shaped transformation failure, which is increasingly the same thing.

Arup (2025). Finance staff approved a \$25.6M transfer after a video call with what they believed was the CFO. The CFO was a deepfake. Cost: direct loss, plus the realization that the same generative capability you are deploying internally is now production-grade in the hands of attackers.

These are the visible failures. The Decay Tax is mostly invisible: the customer who quietly stops using your AI feature, the analyst who stops trusting the dashboard, the support team that builds workarounds because the model "got worse." The McKinsey 2026 AI Trust survey reports 51% of organizations have logged at least one negative AI incident in the past year. The other 49% are not better behaved. They are less instrumented.

The headline metric: AI Half-Life

Every AI system has a **half-life** : the period after which 50% of its accuracy, utility, or user trust has eroded.

Most enterprises cannot tell you theirs.

Mine the number. Then publish it. Then defend it.

This is the metric the next decade of AI ops will be built around. Capability gets benchmarks; decay needs a number too. "What is your AI's half-life?" is the question every board AI committee should be asking by Q4.

A maturity model for resisting decay

Four stages. Be honest about which one you are at.

Stage 1, Blind. You shipped AI. You measure capability and cost. You do not measure decay. You will discover the Decay Tax through a customer escalation or a viral failure.

Stage 2, Aware. You have a postmortem culture. You know drift exists. You have not yet instrumented it. Your AI council talks about it; nobody owns it.

Stage 3, Instrumented. You measure half-life. Eval sets get refreshed on a cadence. You catch silent regression on vendor model swaps within hours, not quarters. You have a number to show the board.

Stage 4, Self-healing. Your systems re-evaluate continuously, route around degradation, alert before users notice. Decay becomes a managed cost, not a hidden tax.

By 2027, every Fortune 500 board will be asked which stage they are at. Most will discover they are at Stage 1.

What to do this quarter

If this frame lands for you, the cheapest first move is not a tool. It is a question.

In your next AI council, ask:

For each AI system in production: what is its half-life, who owns measuring it, and when did we last refresh the eval?

If nobody can answer for any system, you have just found your Decay Tax. The next move is to give it an owner, a number, and a cadence. Tools come after.

The capability era of enterprise AI is roughly over. Every vendor sells capability now. The reliability era is starting, and the enterprises that win it will be the ones that named the problem first.

The Decay Tax has a name now.

The bill is already in the mail.

Shared privately. Please do not redistribute.