

The Case for Runtime Behavioral Attestation

Deployed AI is governed on the wrong clock. The argument, the goal, and the five steps that get there — each step has its own paper.

The argument

Every assurance discipline the enterprise trusts — financial audit, SOC 2, ISO — is **periodic**: the period closes, then someone examines it. Deployed AI does not work in periods. Prompts are edited, models update, tools are added, data shifts — behavior drifts between any two checkpoints. The market has observability (the deployer's own stream), process certification (the management system), and one-off evals (a moment). It does not have a **conflict-free measurement layer a third party can verify**. That missing layer is runtime behavioral attestation.

The goal

Make “is my AI behaving?” into **evidence a third party can accept without trusting the deployer** — and keep that evidence current as the deployment changes. When behavior is measured, recovery is evidenced, and the record is verifiable, AI risk stops being a debate and becomes something an auditor can examine, a board can govern, and — eventually — an insurer can price.

How we get there — five steps, one record

1	BORROW THE TRUSTED DISCIPLINE <i>The audit profession already solved “believable to strangers.”</i> For nearly a century, independent audits against published standards have made financial statements believable to strangers — reasonable assurance, never a guarantee. We apply that exact discipline to AI behavior: published instrument, independent examination, an opinion with stated limits.	<i>Paper: CPA Audit Comparison</i>
2	MOVE IT TO THE RIGHT CLOCK <i>One word changes: periodic → runtime.</i> Every audit you trust examines a closed period. Deployed AI doesn't work in periods — prompts are edited, models update, tools get added, data shifts between any two checkpoints. The blind spot of every trusted audit sits exactly where AI risk lives.	<i>Paper: Runtime Audit</i>
3	DEFINE WHAT YOU MEASURE <i>“99.9% reliable” means nothing for AI.</i> Classical reliability assumes deterministic behavior. For AI, reliability is measured behavioral stability: consistency across dimensions, stability between runs, predictability against declared bounds — every measurement verifiable against a locked, preregistered instrument.	<i>Paper: Behavioral Reliability</i>
4	PROVE THE RECOVERY, NOT PREVENTION <i>A non-deterministic system will sometimes get it wrong.</i> It gets it wrong while every infrastructure light stays green. Resiliency is what you measure instead of prevention: detect degradation early, bound its severity through a defined escalation path, and evidence the return to bounds.	<i>Paper: Behavioral Resiliency</i>
5	MAKE THE RISK PRICEABLE <i>Measurable risk is insurable risk.</i> Fire, telematics, cyber — each time, insurance demand created the measurement standard. AI liability is at the start of that curve, and the measurement record has to exist before anyone can price the risk. We build the record; we do not set premiums.	<i>Paper: Insurability</i>

One record underneath all five

Every step rests on the same record: a **locked, preregistered 120-question instrument** run against the live AI — on demand, on a compliance schedule, or via SDK on a production cadence — with every result **SHA-256 hash-chained** and checkable at a **public verify URL** by a board, auditor, or partner who never has to trust our dashboard. Validation limits are published, and gates that fail are reported as FAIL. Ask for the live demo on the model of your choice.