WHITE PAPER

Independent Assurance

How an Enterprise Can Stand Behind an Agent's Work

Srinivasarengan Varadharajan

Founder & CEO, KTER · 31 October 2026 · kter.ai

ABSTRACT

An enterprise can stand behind an AI agent's work only when three things are separated: the agent that produces the work, the check that clears it, and the person who decides on it. The five earlier KTER papers named the gaps that open when they are not: accountability, blast radius, consequence, visibility and evidence. This paper states the thesis those papers were circling, describes the method KTER built to it, Independent Assurance, and reports what building and running it taught us.

The method is simple to state. Every piece of work an agent delivers is checked in full by a function separate from the agent, and ends in one of three documents. A named person decides. The decision and the evidence are sealed in a record anyone can verify. If the work is wrong, it is done again at no charge. The paper does not claim this removes error. It claims that error becomes visible, attributable and recoverable, which is what accountability has always meant in professional work.

1. The Five Gaps

The earlier papers in this series each named one way an enterprise ends up carrying risk it cannot see. Read together they describe a single structure that is missing, not five separate problems.

Paper	The gap	What it established
The Accountability Gap (July 2026) [1]	Who answers for the output	Every major vendor disclaims liability at the handoff. Accountability closes by contract and operations, not by a feature: named responsibility, agreed scope, a consequence for a missed standard, settled ownership of what the system learns, a person to call, and a reconstructable record.
Trust Has a Blast Radius (July 2026) [2]	What the agent can reach	A permission that is not enforced is a suggestion. The contractual boundary and the technical boundary must both hold, and both must be decided before the agent runs.
The Consequence Gap (July 2026) [3]	What the agent would never do	A person carries an absorbed sense of consequence. An agent has none unless it is specified in advance, in writing, by someone who thought to.
The Visibility Debt (August 2026) [4]	Whether anyone understands it	A green dashboard verifies activity, not judgement. Understanding has to be checked on a schedule by a named person, independent of whether anything has broken.
The Evidence Gap (October 2026) [5]	Whether the claim of success can be relied on	When the agent writes both the result and the report, a reviewer may be checking a claim against another version of the same claim. Acceptance needs a record the agent cannot rewrite, a verifier outside the producer, and a named person with the authority to refuse.

Each paper ended at the same place: the fix is structural, and someone has to build it. This paper is the account of building it.

2. Thesis: Three Separations and One Consequence

An AI workforce is accountable when the work, the check and the decision are held by three different parties, and when a wrong result has a consequence the provider bears. Nothing less closes the five gaps, and nothing more is needed to close them.

Separation 1: the producer never clears its own work. The agent that does the work has no write access to the evidence used to judge it and no say in the acceptance rule. This is the verification boundary the Evidence Gap proposed. It closes the evidence gap directly, and the visibility debt indirectly, because the check produces a readable account of why the work passed or failed.

Separation 2: the check never decides. Assurance clears or returns work; it does not approve it. A named person in the client's organisation decides what happens next, with the evidence in front of them. This is the escalation path from the Accountability Gap and the named reviewer from the Visibility Debt, made structural rather than procedural.

The scope the agent works within is set by that person, in writing, before the agent starts, which closes the blast radius and the consequence gap in the only way they can be closed: by a person specifying, in advance, what the agent may do.

Separation 3: the decision never rewrites the record. Every result, every check and every decision is sealed as it happens, with a reference the client can verify later. The person who decides cannot edit the evidence afterwards, and neither can the provider. This is the reconstructable record from the first paper and the record the agent cannot rewrite from the fifth.

The consequence: the provider answers for the result. Work that was cleared, decided and sealed, and later proves wrong, is done again at no charge, and the correction is recorded against the original. A standard with no consequence is not a standard. This is the one element no vendor terms of service today contain, and the one that makes the other three worth building.

The thesis has a corollary that shaped everything that follows: the three separations are a property of architecture, not of policy. A policy that says the reviewer is independent can be ignored under deadline pressure. An architecture in which the producer has no path to the evidence cannot be.

3. The Method, as the Client Sees It

KTER implements the thesis as three named layers. Independent Assurance clears the work. Executive Oversight decides. The Record seals it. This section describes what each produces for the client. How each is built is KTER's own.

3.1 Independent Assurance

Every piece of work an agent delivers is checked, in full, by a function that is separate from the agent that produced it. The check reads the work, reads what the work relies on, and ends in exactly one of three documents.

Document	What it means	What happens next
Certificate	Every statement in the work met its standard	The work goes to the person who decides
Work Notice	The work has findings, each stated as what failed, where, and why	The person decides whether to send it back with their own input
Declination	The work was outside the agent's scope and was not attempted	The person decides whether the scope should change

Nothing leaves assurance without one of the three. A Work Notice is not a failure of the system; it is the system doing what a reviewer is for. A client who never sees one should ask why.

3.2 Executive Oversight

A named person in the client's organisation sees the document and the evidence behind it before anything is acted on. They approve, hold, or send the work back with their own input. Sent-back work is corrected against the findings they saw and checked again in full, and every version is kept and linked. The scope an agent works within is set by that person, in writing, before the agent starts. Autonomy is granted by that person in steps, and is visible on the agent's own card as a record of trust earned.

3.3 The Record

Each document is sealed when it is issued, with a reference the client, a counterparty, an auditor or a regulator can verify independently: that the document exists, when it was issued, and that it has not changed since. A client may publish its record of certificates. Work Notices and Declinations are never public. The reference is the answer to the first paper's last question: what happened, and why, can always be reconstructed.

3.4 The consequence

Work that carries a Certificate and later proves wrong is redone at no charge, and the correction is recorded against the original reference. That is the whole of the commitment. It is deliberately narrow, because a commitment that can be met is worth more than one that cannot.

4. What Running It Taught Us

The method in section 3 is not the method we started with. Three lessons from KTER's internal pilot work in September and October 2026 changed it. None involves client material, and none is specific to any one vertical.

Honesty has to be rewarded by the structure, not just permitted. The Evidence Gap argued for an honest blocked state. Building one is harder than allowing one: an early version of assurance returned work for the very sentences in which the agent said what it could not do. An agent that says what it could not do has done its job, and the check now treats it that way. Any assurance system that punishes a declared limitation will train its agents to stop declaring them.

Stricter than a careful person is not safer. An early version of assurance was stricter than a careful human reader and returned work whose sources were correct. A check that is wrong in the strict direction is still wrong, and it costs the client the same correction cycle. The standard now is the one a careful professional would apply: does the source say the thing.

The first-pass rate is low, and that is the point. In the internal pilot, most work did not certify on its first pass. The alternative to a Work Notice is not a Certificate; it is the same work delivered as complete and found later, or not found. The honest number is reported in place of a flattering one, and it will move as the agents' calibration improves. We would be wary of a vendor whose first-pass rate is close to a hundred percent, for the reason section 5 gives.

5. Practice: How an Enterprise Uses It

The method is only useful if it fits into how an organisation already works. Four practices make it fit.

Match the depth of oversight to the consequence. The Evidence Gap proposed four consequence tiers. In practice they map onto who decides, not onto how much is checked: assurance checks every piece of work the same way, and the tier decides who must see the result before it is acted on.

Consequence	Example	Who decides
Low	A draft for internal discussion	The person who asked for it
Moderate	Analysis others will rely on	A named reviewer
High	A change to a customer or supplier record	A named reviewer, and a named owner of the correction if it is wrong
Critical	A regulatory filing or a payment instruction	The accountable officer, named in advance

Grant autonomy in steps, and only forward. A new agent starts with every result decided by a person. As its certification rate on a class of work holds, the person may let that class proceed on a Certificate alone, with the decision recorded as a standing one. A Work Notice on that class returns it to full oversight. Autonomy is a record of trust earned, visible on the agent's own card.

Run the failure test before going live. Before any workflow is trusted, make a required source unavailable in a test setting and ask the agent to proceed. A passing workflow reports the source as unavailable, holds the task, and cannot mark it complete. Then restore the source, repeat, and confirm the evidence belongs to the new run. Repeat the test whenever the model, the tools or the acceptance rule changes. This is the Evidence Gap's pilot test, unchanged, and it is the single most informative hour an organisation can spend on an AI vendor.

Ask for the reference, not the report. A completion report is a claim. A reference is a sealed record anyone can verify. When an agent's work is handed to a colleague, a client or a regulator, the thing handed over is the reference. Any vendor whose output cannot be verified by someone outside the vendor has left the evidence gap open, whatever the report looks like.

The buyer's questions from the earlier papers still apply and are not repeated here. One is added: **show me a Work Notice.** A vendor that can only show certificates has either a flawless agent or no independent check. The second is more common.

6. Limits and Open Questions

The method makes error visible, attributable and recoverable. It does not make error rare, and it does not make the reviewers right.

The check can be wrong. An independent check is stronger evidence than the agent's own report, not proof. A statement the check accepts for a bad reason passes. The failure test and the named person are the defences, and neither is complete.

A source can be true and stale. Assurance checks that a source says the thing. It does not check that the thing is still the case. Dated sources are the client's standard to set, and the method records the date so the question can be asked.

Judgement is cleared, not evaluated. Assurance establishes that a judgement is presented as one and rests on what it cites. Whether it is a good judgement is the deciding person's question, and the method puts the evidence in front of them rather than answering it for them.

The consequence is narrow by design. Redoing the work at no charge is the commitment. Consequential loss from acting on wrong work is a question for the contract and the law of the place where the work is consumed, and is being defined with counsel; the method does not pretend to settle it.

The rate will be contested. A low first-pass certification rate will be read by some as proof the agents are weak and by others as proof the check is strict. Both readings have a point. The honest number is better than the flattering one, and it will be published as it moves.

The structure cannot supply what the third paper said no document can. An agent does not carry a person's absorbed sense of consequence. A scope set by a person and a person at every decision are the best available substitute, and they are a substitute.

Closing: The Structure Is the Product

The first paper in this series said the accountability gap closes by structure, not by a feature, and that most AI companies would not build the structure because it is operations work. This paper is the record of building it anyway. Three separations, one consequence, three documents, one reference. None of it is clever. All of it is deliberate.

The test an enterprise should apply is the one the fifth paper ended on, now with the answer attached. Before accepting "done", ask to see the run. With Independent Assurance the answer is a reference, and the reference is enough.

SOURCES

[1] Varadharajan, S. The Accountability Gap: Why Every Enterprise AI Deployment Has a Hidden Liability Problem. KTER, 16 July 2026.

[2] Varadharajan, S. Trust Has a Blast Radius: Why Boundaries Have to Exist Before an Agent Runs, Not After. KTER, 23 July 2026.

[3] Varadharajan, S. The Consequence Gap: Why an Agent Can't Restrain Itself the Way a Person Does. KTER, 27 July 2026.

[4] Varadharajan, S. The Visibility Debt: Why Running an Agent and Understanding an Agent Are Not the Same Thing. KTER, August 2026.

[5] Varadharajan, S. The Evidence Gap: What an Enterprise Needs Before It Can Accept an Agent's Claim of Success. KTER, October 2026.

The lessons in section 4 are from KTER's internal pilot work, September and October 2026, with no client material. The consequence tiers and the autonomy practice are the author's proposals.