

The Handoff With No Receiver

Buy-side probe into an unowned assurance gap · August 2026

AI assurance measures the trustworthiness of AI systems, and presumes the assessor can reach the system. Safeguards frameworks specify what a document must contain and what standard it must meet, and have no field for how it was produced. Both scopes are correctly drawn. Nothing sits between them. This is what six practitioners said when asked what control is actually operating.

6 substantive replies from named practitioners	44h elapsed between first send and sixth reply	2 sides confirming the same conclusion independently	0 mechanisms found that own the handoff
---	---	---	--

What the replies established

Finding	What it changes
Adoption has already happened	A World Bank ESF reviewer put it at almost every impact assessment product now crossing his desk. This is not a forecasting problem or a policy-window problem. The documents are already in the system.
Polish has stopped being a signal	The profession read fluency as a proxy for diligence for decades without ever naming it as a control. That proxy now carries no information, and nothing has replaced it.
The operating control is relational	Experienced reviewers can often tell, because they know the author's prior work and the tone is wrong. Real and skilled — and undocumented, unmeasurable, non-transferable, and expected by its own practitioners to fail as models improve.
The lever is capacity building	Named unprompted from the buy side. This inverts the commercial sequence: the institutional entry point is training reviewers, not selling them assurance.
Detection is the wrong target	Confirmed independently from the technical side and the sector side. Where evidence originates as testimony, any check runs on a translation and returns the translating model's fingerprints. There is nothing to recompute.

Where the gap sits



The scope that falls out of it

Assurance over how a document was produced, for documents whose ground truth cannot be recomputed.

It sits between two scopes rather than inside either. Not AI assurance, which stops at the system boundary. Not domain compliance, which tests the artefact against a standard. The transition between them.

It is defined by the non-recomputable case. Where a value can be regenerated and diffed, existing approaches already reach it and there is no shortage of people building them. Where the origin was speech and the destination is a compression, there is nothing to recompute, and the only available control is a record made while the work happened.

Its instrument is disclosure and attestation, not detection. What was model-assisted, what it was checked against, who checked it, and what they found.

Why detection cannot be the answer

Consider one chain in social performance work. A person speaks in a community meeting, in their own language, with hesitation and qualification. That becomes a fieldworker's note, then a coded entry in a tracker, then a section in a report, then a summary of that report, then a dashboard indicator, then a funder decision, then a resource allocation that lands back on the same person.

Every arrow is a compression, and several are now automated. The losses are systematic rather than random: each step keeps what its schema can hold and drops the rest. The qualification goes. The dissenting household goes. The case that did not fit a category goes. There is no original to diff the output against, because the original was spoken. What can be recorded is the production history — and it has to be recorded while the work happens, because it cannot be reconstructed afterwards.

The procurement-legible name for this, for a budget line rather than a brand, is **Evidence Production Assurance**. Deliberately dull, and held loosely until a client uses a phrase of their own. Probe respondents are not named here; the underlying correspondence is held and quoted only with permission.

Liezl Coetzee · Sociable Systems · sociable.systems · liezlc@gmail.com