HER$_2$™ — Forensic Evidence System for Artificial Intelligence

Recomputable, non-normative evidence for AI interaction behavior

————

The Problem

When AI systems are audited, investigated or litigated, dashboards, logs, and narratives collapse.
- Vendor logs are mutable
- Explanations are post-hoc
- Model behavior cannot be independently reconstructed

In legal disputes, claims per se are not reliable evidence.

————

What HER$_2$™ Is

HER$_2$™ is the forensic evidence system for AI.

It produces recomputable, cryptographically verifiable records of how the AI system behaved — over time — under declared conditions.

HER$_2$™ preserves the facts of what happened, in what order, and with which evaluators, without modifying the model/s.

————

What HER$_2$™ Produces
- Interaction Integrity Records
  Structured JSON artifacts binding inputs, outputs, context, and provenance.
- Temporal Causality Records
  Append-only lineage across time (drift, divergence, recovery preserved as evidence).
- Cross-Evaluator Evidence
  Parallel records across evaluators without forced consensus.
- Signed Evidence Bundles
  Manifest, hashes, signatures, and recomputation instructions.

All artifacts are portable, inspectable, and independently cryptographically verifiable.

————

What HER$_2$™ Is Not
- Not a certification or compliance judgment
- Not an enforcement or policy engine
- Not a model controller or modifier
- Not a dashboard or explanation layer

HER$_2$™ does not adjudicate facts — it preserves evidence.

————

Who Uses HER$_2$™
- Venture Capital / Private Equity
  Verify AI behavior during diligence — and prove it later.
- Banks, Fintechs, Insurers
  Reconstruct AI decisions under audit or dispute.

- Audit & Compliance Firms
  Independently replay model behavior beyond control reviews.
- Litigation & Expert Witnesses
  Produce neutral, court-inspectable AI evidence.
- Insurance Underwriters
  Reconstruct incidents to price and assess AI risk.

———

Deployment
- Runs offline, in CI, or inside agent runtimes
- No access to model weights or training data required
- Vendor-agnostic by design

———

Status

Patent pending · Private pilot licenses available

HER$_2$™ is derived from the public HER2NI research protocol family and delivered as a private forensic evidence system.
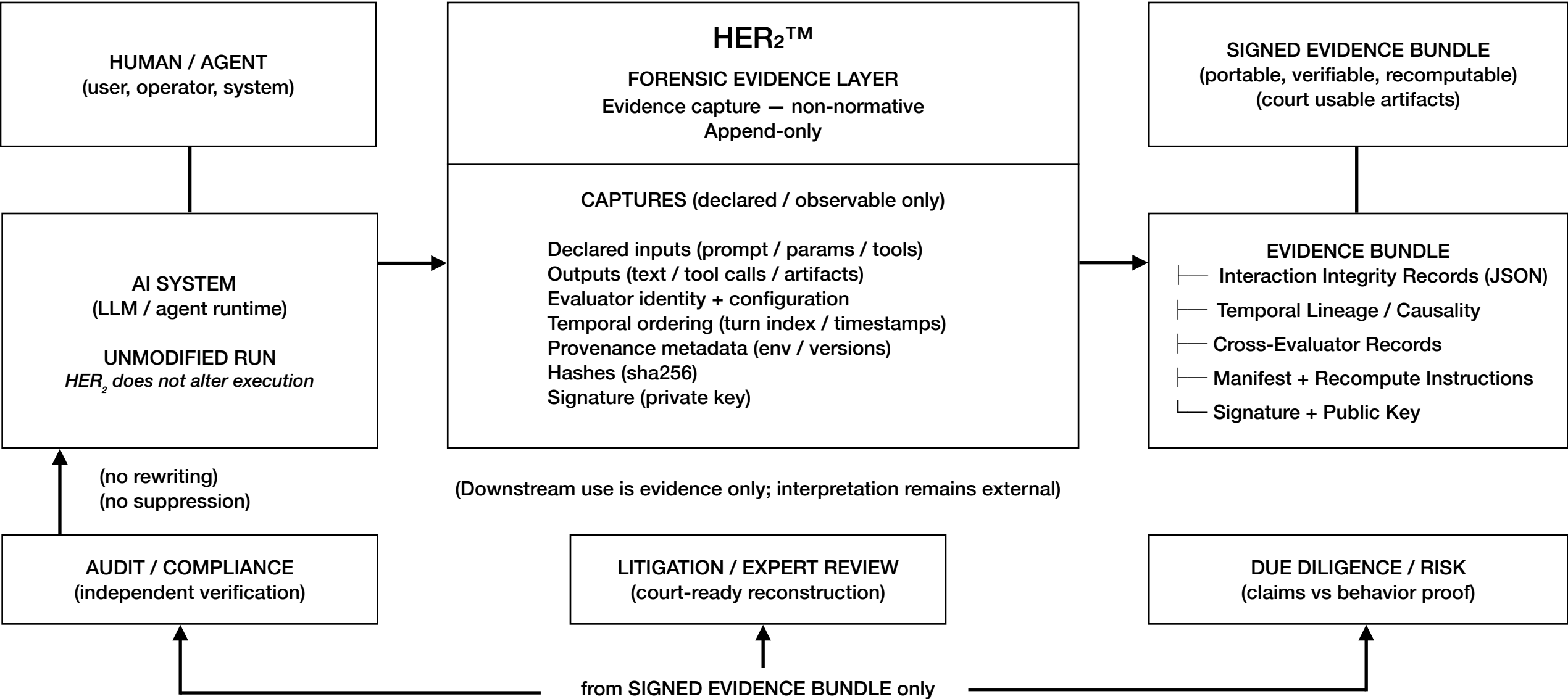
———

Summary

HER$_2$™ is not a model
HER$_2$™ does not judge correctness
HER$_2$™ produces court-usable artifacts
HER$_2$™ survives hostile review

When narratives collapse, only recomputable records remain.
HER$_2$™ is your evidence layer.

———

Annexed

FIG. X — HER$_2$™ Architecture & Interaction Evidence Flow (Model-Agnostic)

FIG. X — HER₂™ Architecture & Interaction Evidence Flow (Model-Agnostic)

HUMAN / AGENT
(user, operator, system)

AI SYSTEM
(LLM / agent runtime)

UNMODIFIED RUN
HER₂ does not alter execution

(no rewriting)
(no suppression)

HER₂™
FORENSIC EVIDENCE LAYER
Evidence capture — non-normative
Append-only

CAPTURES (declared / observable only)

Declared inputs (prompt / params / tools)
Outputs (text / tool calls / artifacts)
Evaluator identity + configuration
Temporal ordering (turn index / timestamps)
Provenance metadata (env / versions)
Hashes (sha256)
Signature (private key)

(Downstream use is evidence only; interpretation remains external)

SIGNED EVIDENCE BUNDLE
(portable, verifiable, recomputable)
(court usable artifacts)

EVIDENCE BUNDLE
├── Interaction Integrity Records (JSON)
├── Temporal Lineage / Causality
├── Cross-Evaluator Records
├── Manifest + Recompute Instructions
└── Signature + Public Key

AUDIT / COMPLIANCE
(independent verification)

LITIGATION / EXPERT REVIEW
(court-ready reconstruction)

DUE DILIGENCE / RISK
(claims vs behavior proof)

from SIGNED EVIDENCE BUNDLE only

CAPTION:
HER₂™ records interaction behavior as evidence, independent of model internals, explanations, or operator custody.

© STEVE BLACK (IE) | Tbilisi, Sakartvelo