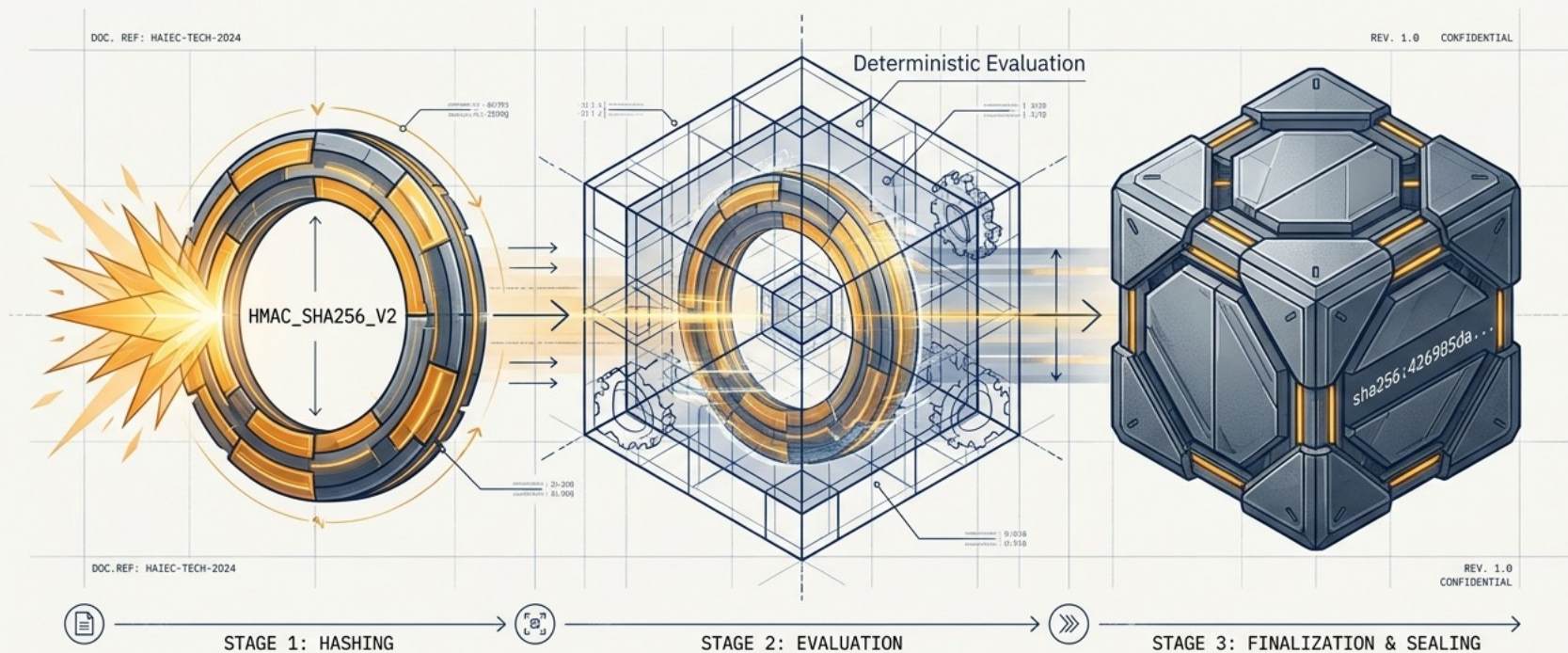


Evidence-Bound Assurance for Consequential AI

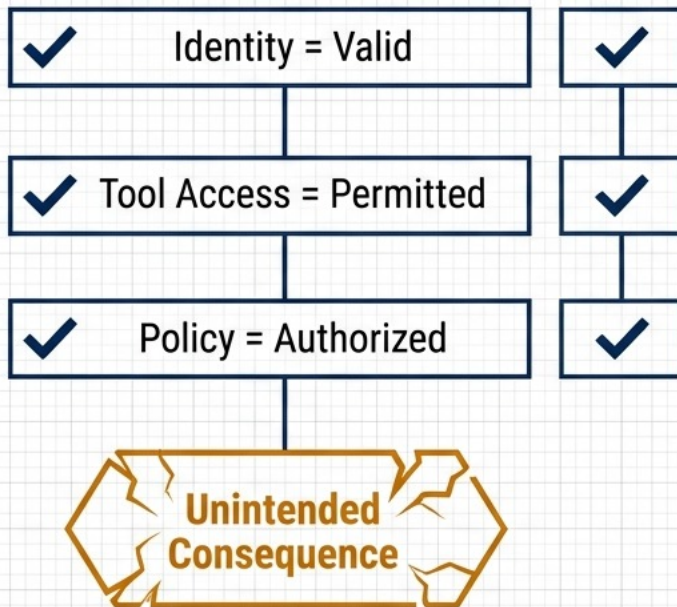
The Cryptographic Mechanics of HAIEC



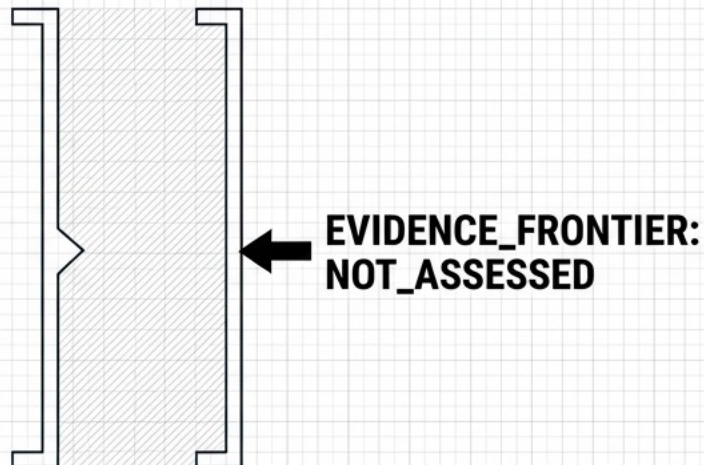
Permission is Not Delegation

A valid tool, called by a **valid identity**, can still cause an unintended **consequence**. How do you defend the deployment decision?

Path A: The Traditional Trap



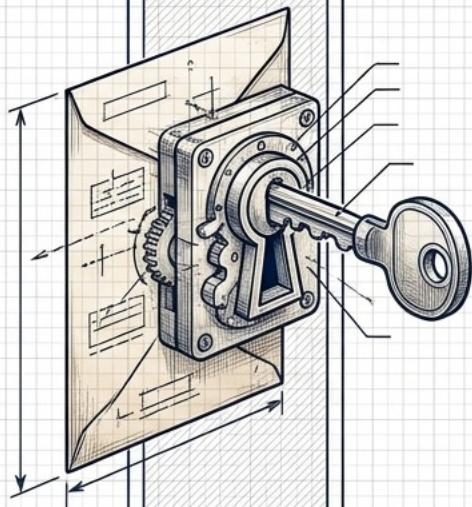
Path B: The HAIEC Defense



The Architecture of Cryptographically Bound Assurance

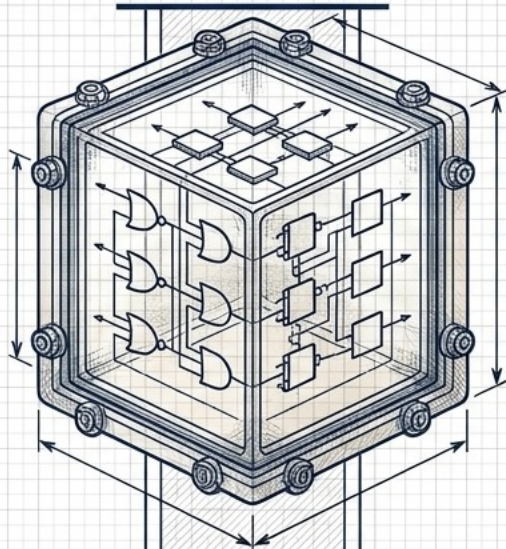
The Verifiable Chain of Custody

Input Integrity



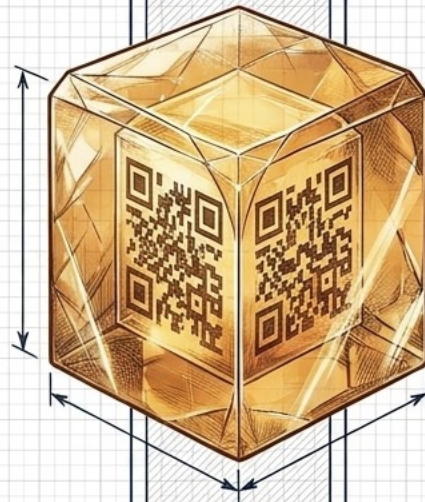
**V2 Payload-Bound
Attestation**

Processing Integrity



**Deterministic
Evaluation Engine**

Output Integrity

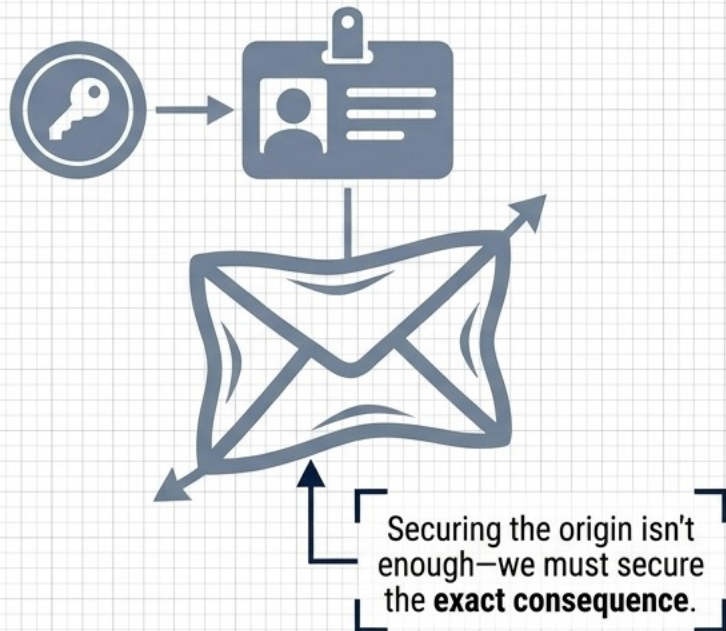


**Deterministic
Integrity Digests**

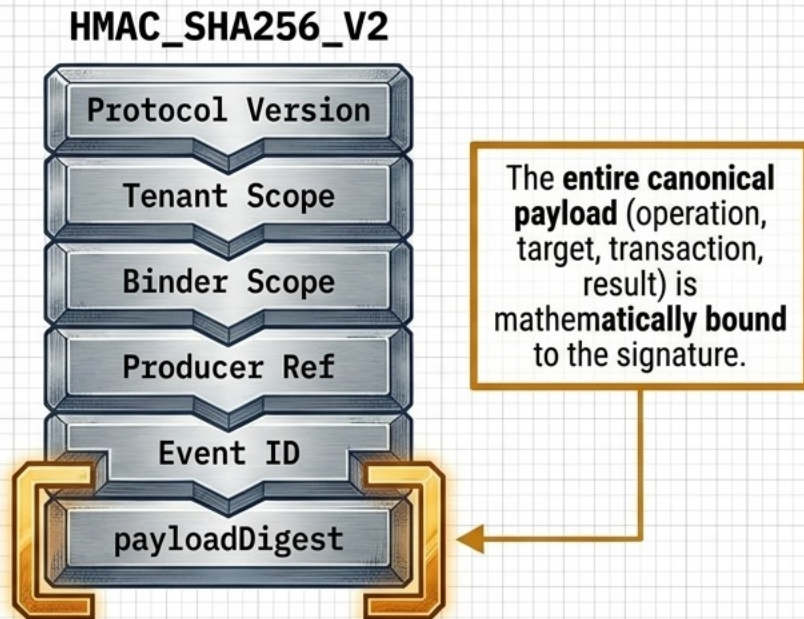
Pillar 1: Input Integrity

Moving Beyond Source Identity to Consequence Locking

The Gap



The Solution



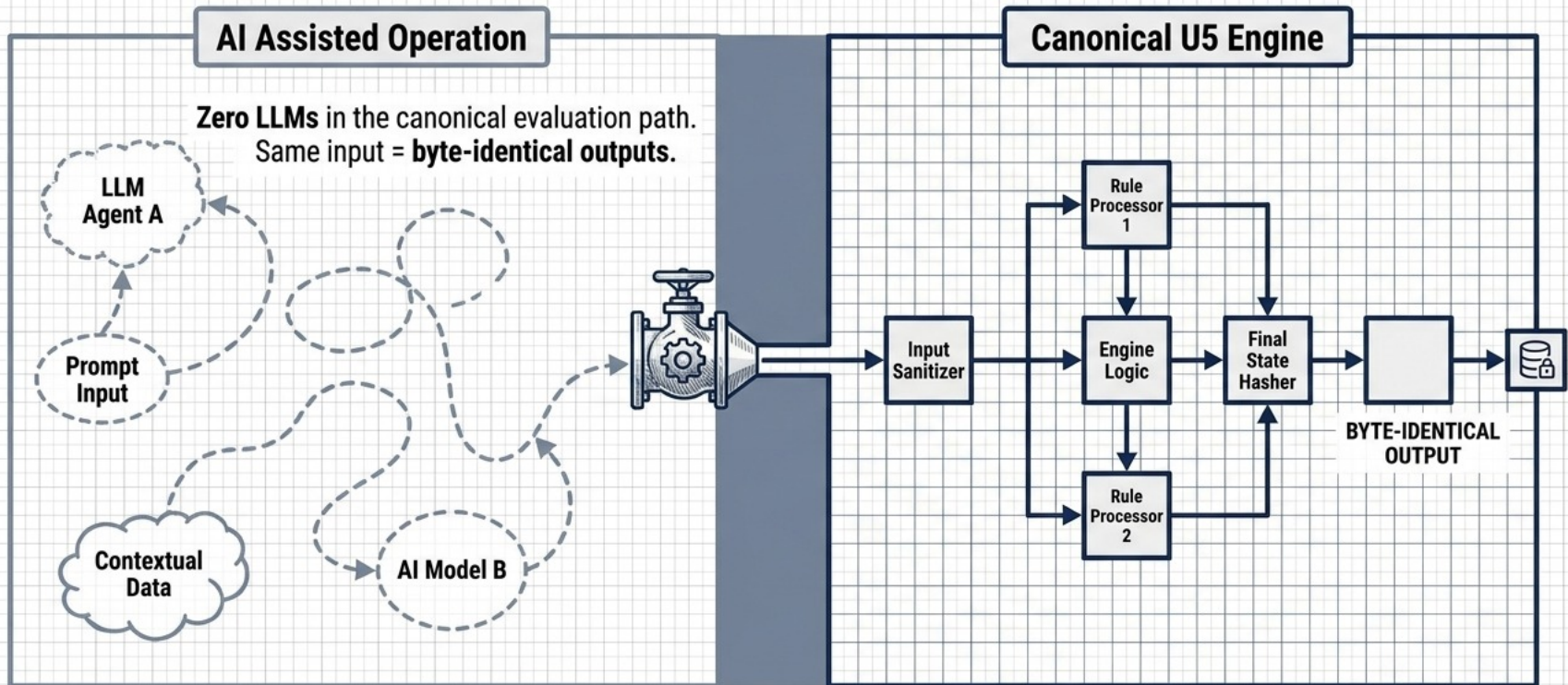
Why V2 is Mandatory for Independent-Effect Proof

Only V2 ξ Only V2 secures the entire canonical payload. This makes it tamper-evident and the only protocol eligible to establish an “independent-effect witness” (UCP-08 R01 receipt).

	Legacy V1_IDENTITY	<u>V2_PAYLOAD_BOUND</u>
Identity Bound	Yes	Yes
Payload Bound	No	Yes (via <code>payloadDigest</code>)
Replay Resistance	Weak – payload is malleable	 hash Strong - tampering invalidates hash
Attestation State	AUTHENTICATED	AUTHENTICATED + V2_PAYLOAD_BOUND

Pillar 2: Deterministic Evaluation Processing

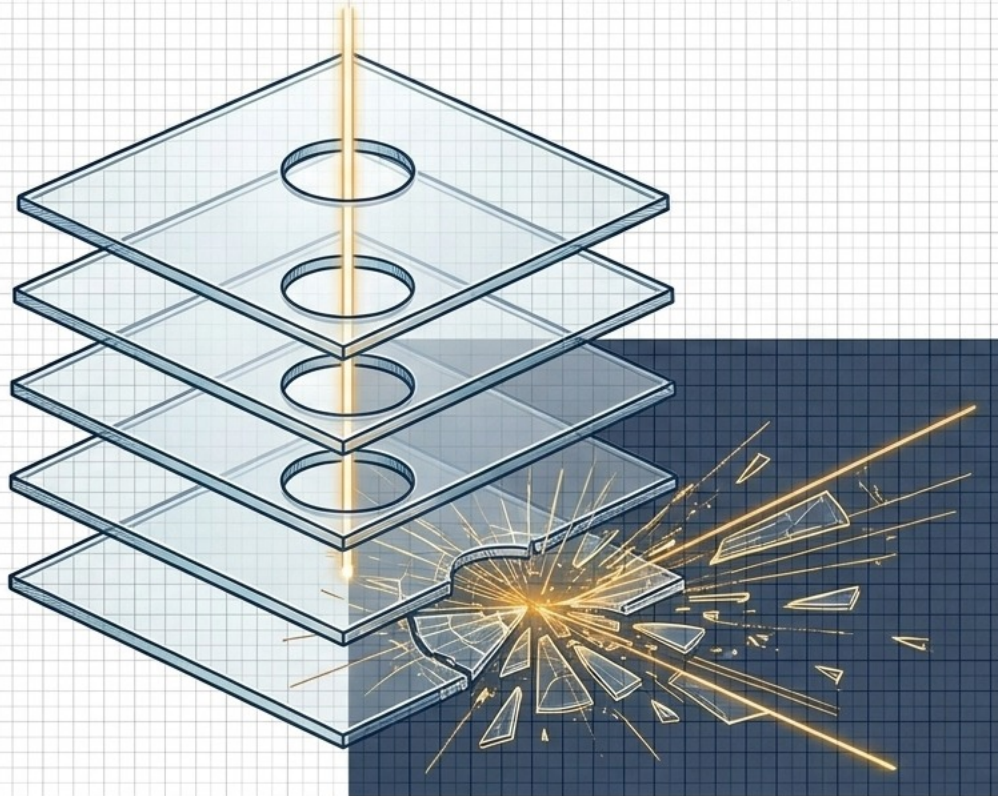
AI Assesses. Algorithms Evaluate.



The Five Independent Evidence Planes

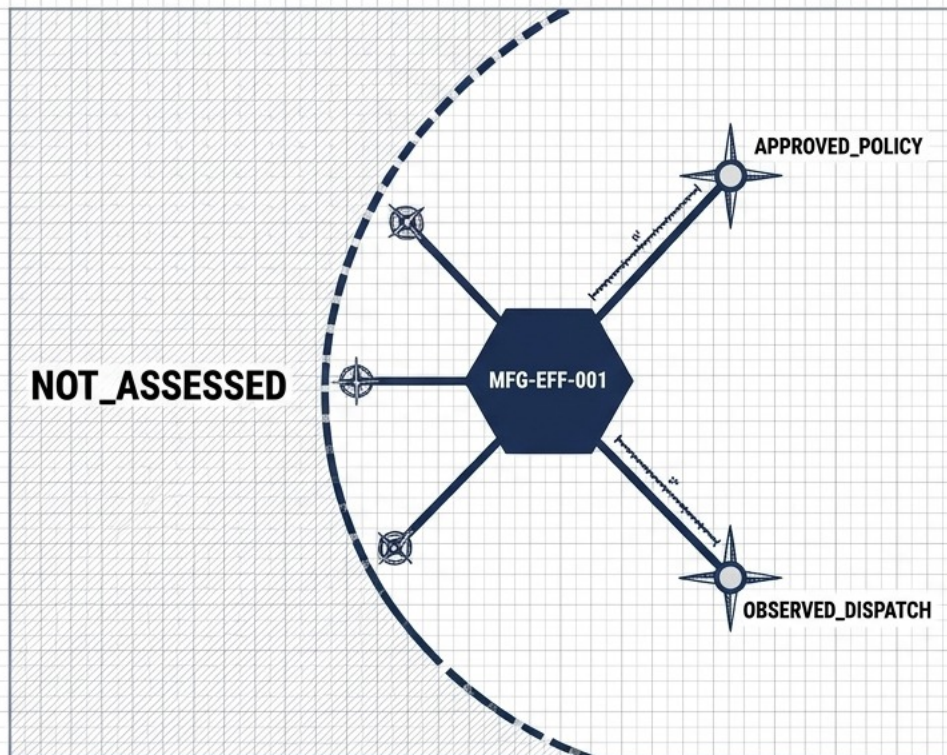
Evidence planes are independent and never merged. Divergence is a factual finding, not an arbitrary verdict.

- 1. REQUESTED**
(What the AI was asked to do)
- 2. POLICY-AUTHORIZED**
(What rules permit)
- 3. EFFECTIVELY GRANTED**
(What credentials allow)
- 4. CODE-CAPABLE**
(What the application can execute)
- 5. OBSERVED**
(What actually happened at runtime)



Rule Adjudication & The Proof Frontier

A rule **MATCH** means a factual divergence was found. If evidence is missing, the proof stops at the frontier. No silent passes.



HAIEC Report - Confidential

4 - Where does the proof stop? - PARTIAL

The proof stops at 5 unresolved evidence-frontier item(s) and 1 open evidence need(s) - listed verbatim.

- 5 evidence-frontier item(s) in state NOT_ASSESSED.
- Producer completion: COMPLETED.
- US reason codes: NOT_EVALUATED, REQUESTED_NOT_AUTHORIZED.
- Not assessed: AI-PROOF-002: File counts are not represented in this artifact
- Not assessed: AI-PROOF-012: Source container continuity not established - repository analysis cannot see build or deployment attestation
- Not assessed: AI-PROOF-014: Exact action execution identity binding not established - source defines code, not the deployed principal
- Not assessed: AI-PROOF-015: IAM authority not established by repository analysis - provider grants are external to source
- Not assessed: Agent Insentery: not established in this evidence basis - agentTopology: producer not connected to this basis
- Not assessed: How Each Agent Operates: not established in this evidence basis - executionArchetype: producer not connected to this basis
- Not assessed: Blast Radius / Shared Access: not established in this evidence basis - agentTopology: producer not connected to this basis; sharedSubstrate: producer not connected to this basis; egress: producer not connected to this basis; resourceScope: producer not connected to this basis
- Not assessed: Human Control and Delegated Authority: not established in this evidence basis - humanControl: producer not connected to this basis; delegatedAuthority: producer not connected to this basis
- Not assessed: Deferred, Persistent and Replayable Actions: not established in this evidence basis - deferredExecution: producer not connected to this basis; actionComposition: producer not connected to this basis; replay: producer not connected to this basis
- Not assessed: Memory, Retrieval and Model-Context Influence: not established in this evidence basis - memoryContext: producer not connected to this basis
- Not assessed: Evaluation Integrity / Self-Interest Surfaces: not established in this evidence basis - evaluationIntegrityAccess: producer not connected to this basis
- Not assessed: Security Posture: not established in this evidence basis - securityScan: producer not connected to this basis; consequenceDelta: producer not connected to this basis
- Not assessed: Sensitive Data and Data Boundary: not established in this evidence basis - memoryContext: producer not connected to this basis
- Not assessed: Recovery, Containment and Reversibility: not established in this evidence basis - deferredExecution: producer not connected to this basis
- Not assessed: Planner, Executor and Objective Continuity: not established in this evidence basis - delegatedAuthority: producer not connected to this basis
- Not assessed: Tool, MCP and Supply-Chain Provenance: not established in this evidence basis - capabilityChain: producer not connected to this basis
- Not assessed: Operational Auditability and Runtime Witness: not established in this evidence basis - provenance: producer not connected to this basis

Next evidence

- Requested capability was not established as authorized by policy evidence.

Pillar 3: Deterministic Integrity Digests

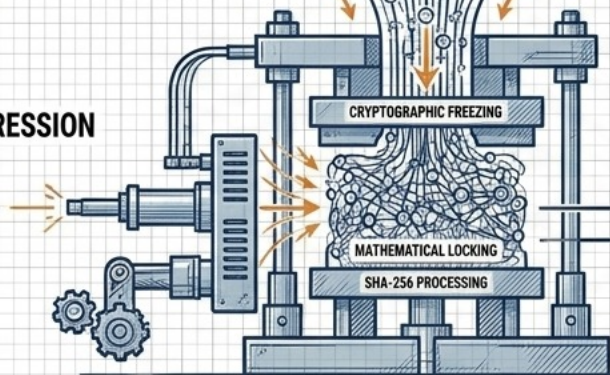
Integrity by Construction, Frozen in Cryptographic Amber

Persisted evaluation snapshots are frozen. They prove **integrity mathematically**, locking the exact state of what was evaluated.

THE EVALUATION



THE MECHANICAL COMPRESSION



THE OUTPUT



Anatomy of a Frozen Evaluation

This is not an assumption. It is an exact projection of persisted evidence, completely divorced from hallucination.

HAIEC Report - Confidential

Master Assurance Report

Exact frozen Evaluation — deterministic presentation. This document is a projection of persisted evidence; it is not the Assurance decision engine and cannot recompute the Evaluation. HTML and PDF share the same truth; presentation differs by design (SAME_TRUTH != IDENTICAL_PRESENTATION).

Evaluation identity

Evaluation	⊗	bce94772-df56-447b-8625-0646915f7e25:assurance:1.1
AI System		SYS-217
Overall Assurance (canonical U5)	⊗	REVIEW
Methodology version		1.1
Time basis	⊗	EVALUATED_SNAPSHOT — current-state context never enters this report
Scenario run(s)		syn-world-sys-217-wrong_target
Core truth digest		0333d6221e27afd8b7d4a567156b8a49f118f305e035a9eda578834c85b21d49
Evidence bundle digest		34e0384eefeddb85b0bed776add8ecace5fe43c18dd3291a04eb95c61af5fd5
Report identity		master-report:bce94772-df56-447b-8625-0646915f7e25:assurance:1.1 : schema master-report-schema-1.0.0 : projection master-report-1.3.0
Presentation profile		BASE (presentation only — excluded from the truth digest)

Unique deterministic run.

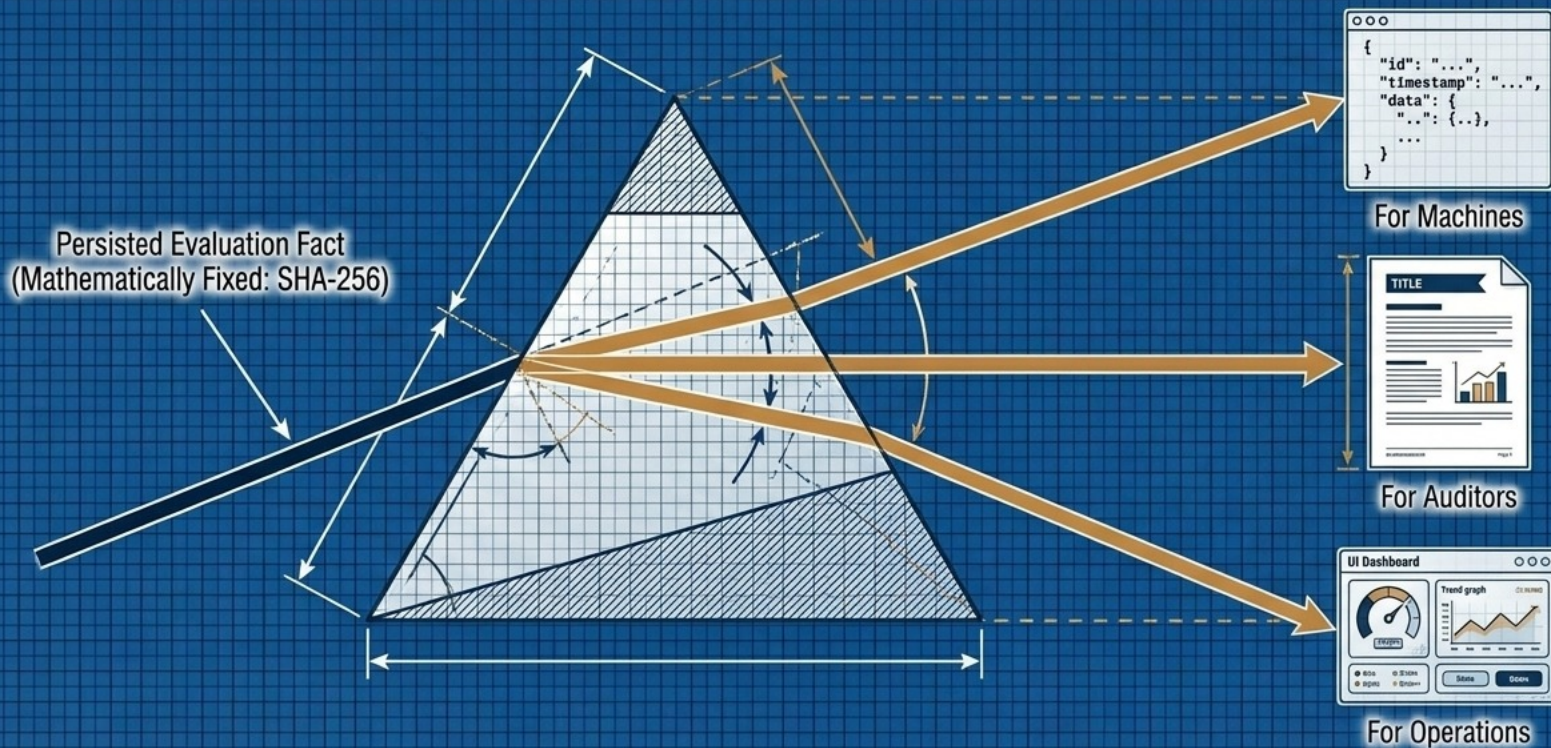
REVIEW disposition.

Cryptographic anchors locking the document.

EVIDENCE_ONLY basis.

SAME_TRUTH != IDENTICAL_PRESENTATION

Presentation differs by design to serve different audiences. The underlying cryptographic truth never deviates by a single byte.

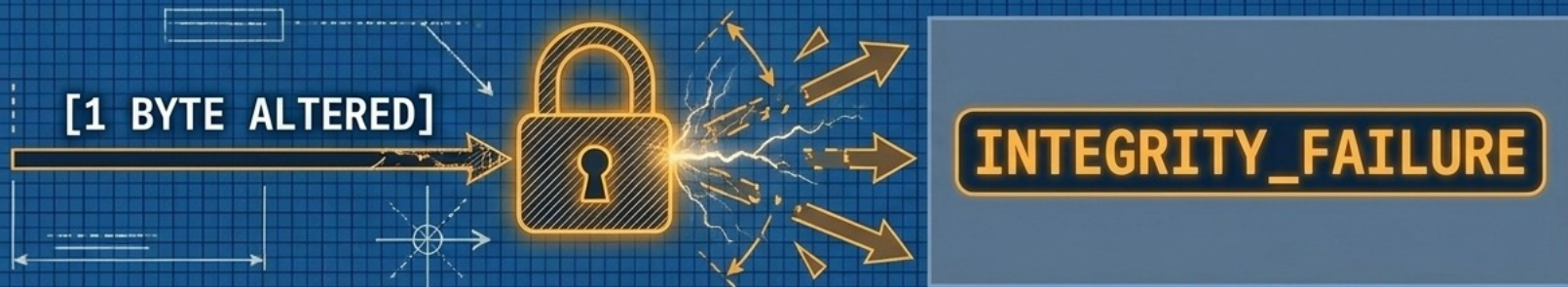


The Fail-Closed Contract

No graceful degradation of truth. If an evaluation is tampered with, unauthorized, or missing evidence, the system fails closed. Absolute blackout.



Digest matches → Content renders flawlessly.



'We Govern Actions' Is Not Enough

Traditional tools establish what is possible and what is permitted. Only HAIEC establishes what was actually caused, backed by independent-effect evidence.



	Scanners / Governance	IAM / Gateways	HAIEC (Evidence-Bound Assurance)
Influence Chain	Pattern matching	None	Tracks influence from prompt to event
Consequence Graph	Policy text only	Permitted access	Traces exact path to actual business consequence
Evidence Frontier	Counts vulnerabilities	Logs authentication	Maps where proof explicitly stops

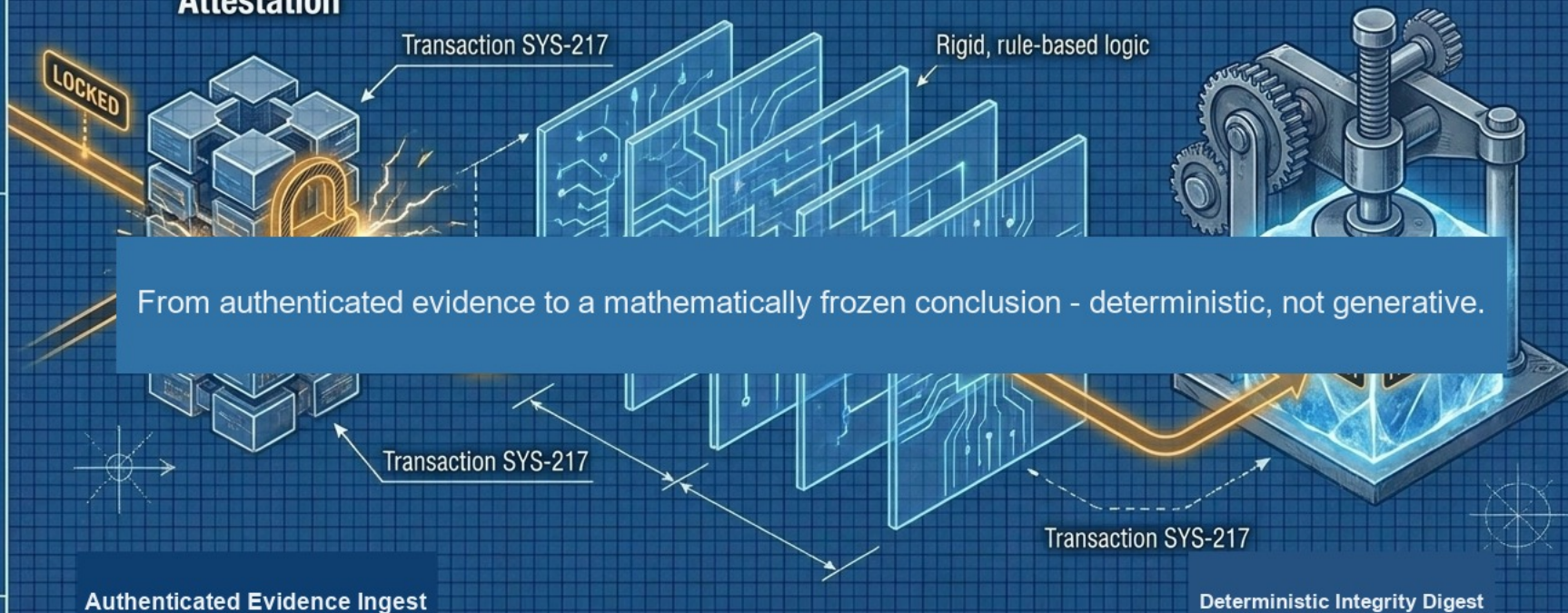
The End-to-End Bounded Path

From cryptographically signed origin to a mathematically frozen conclusion, with zero LLM hallucination.

**V2 Payload-Bound
Attestation**

Deterministic Processing Grid

SHA-256 Truth Digest



Authenticated Evidence Ingest

Deterministic Integrity Digest

Defend Your Deployment

Know exactly what your AI can cause, and know exactly where the proof stops.

Scope an Enterprise Assurance POC

Evaluate one consequential AI system against a defined assurance question. Receive technical findings, inspectable evidence, explicit gaps, and a prioritized action plan.