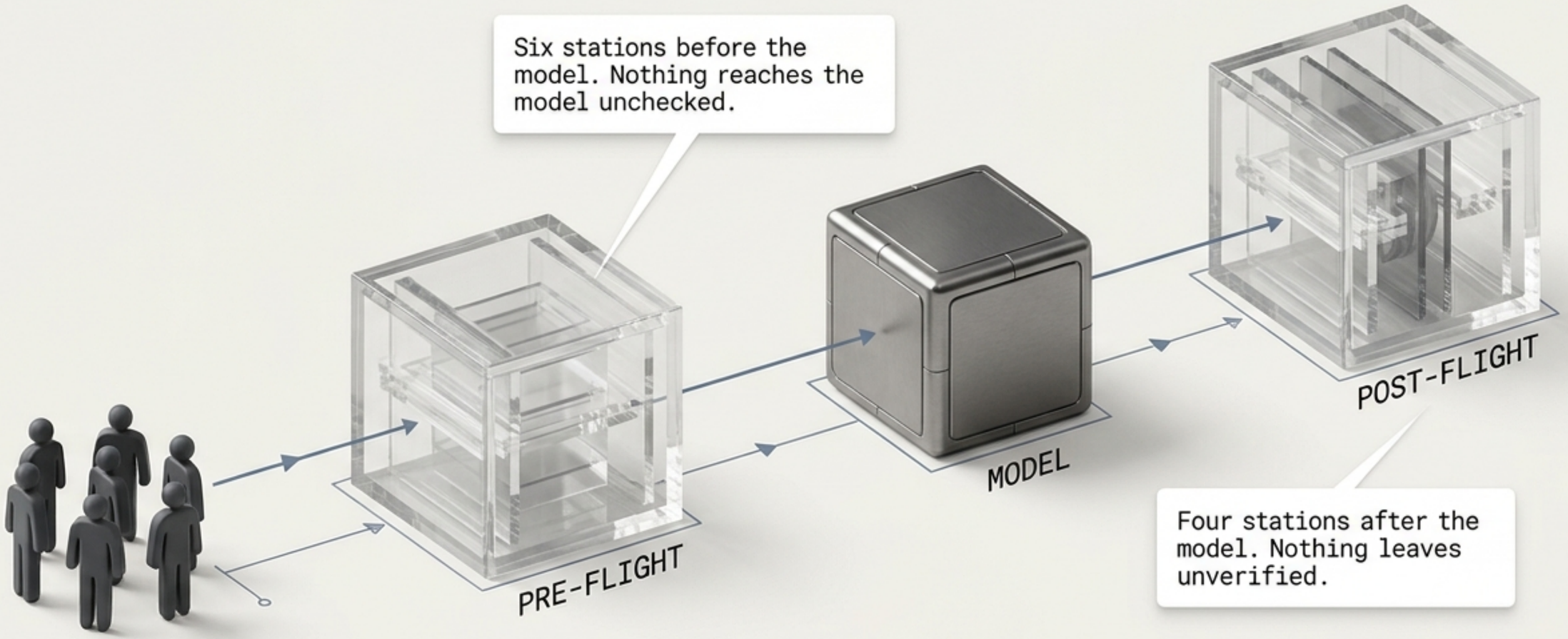
BULWARK: Every AI decision proven before it runs

The cryptographic guardrail for AI governance.

A cryptographic guardrail between your people and your AI models



Every request is inspected, and every verdict is recorded as cryptographic evidence.

Pre-Flight Identity & Data: Validating the source and masking the payload

S01 PRE

Identity & Purpose

Caller authentication, intent classification, prompt-injection detection.

S02 PRE

Data Classification

Sensitivity tagging, token-level redaction before the model sees the payload.

S03 PRE

Personal Data & Secrets

Credential scanning, PII detection, secret masking.

Pre-Flight Security & Spend: Defeating manipulation and gating resources

S04 PRE

Manipulation

Adversarial prompts, jailbreak patterns, social engineering vectors.

S06 PRE

Spend & Rate

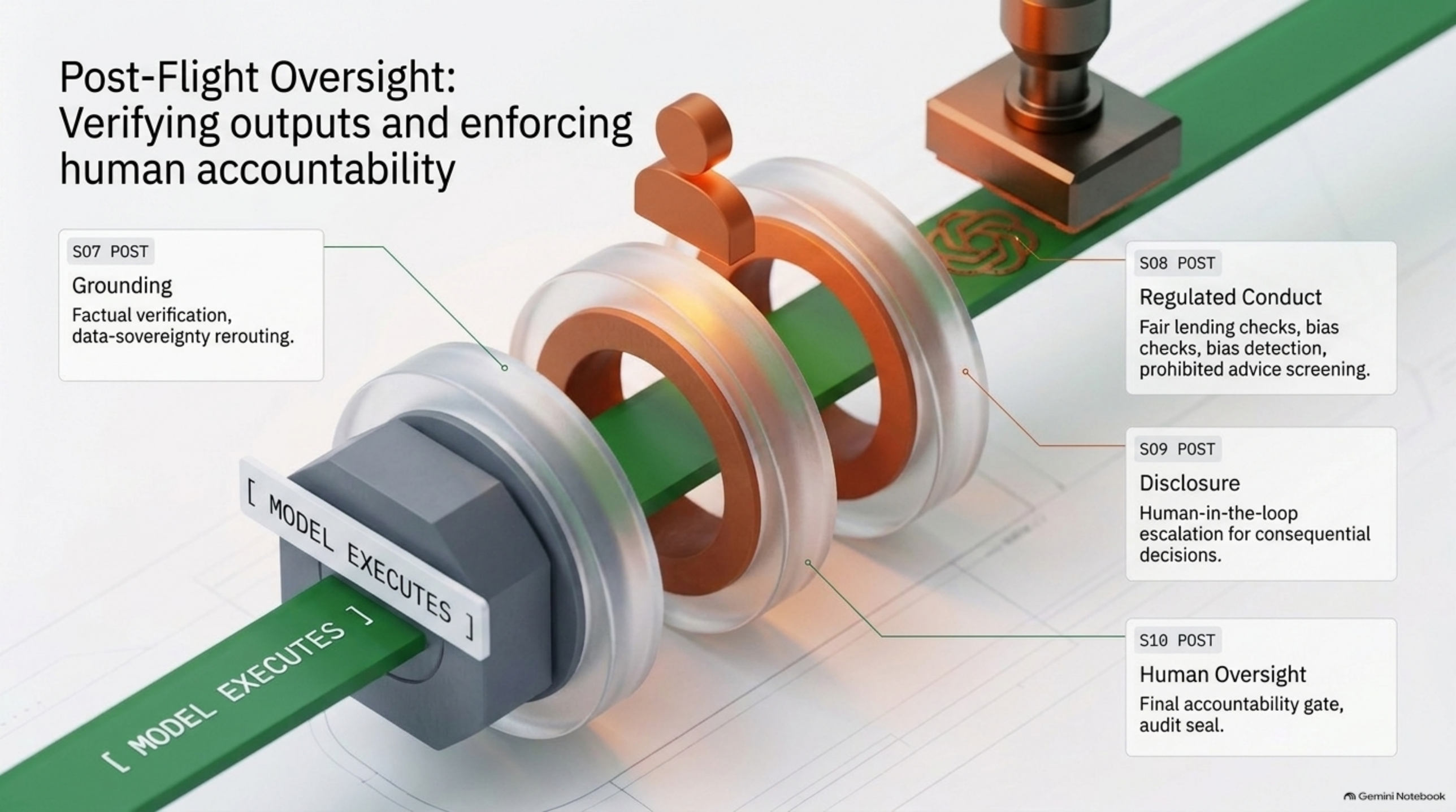
Budget gates, token-rate throttling, cost attribution.

S05 PRE

Model Authorization

Tier-based model access, deployment region enforcement.

Post-Flight Oversight: Verifying outputs and enforcing human accountability



S07 POST

Grounding

Factual verification, data-sovereignty rerouting.

S08 POST

Regulated Conduct

Fair lending checks, bias checks, bias detection, prohibited advice screening.

S09 POST

Disclosure

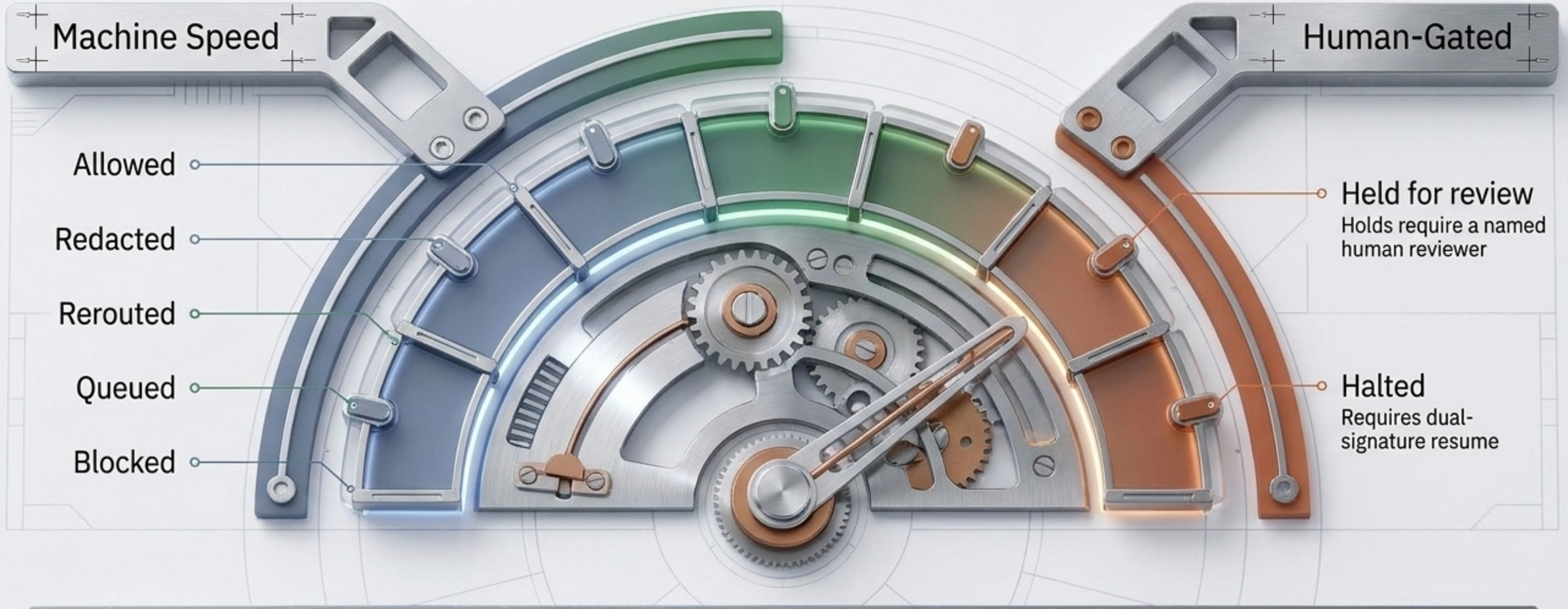
Human-in-the-loop escalation for consequential decisions.

S10 POST

Human Oversight

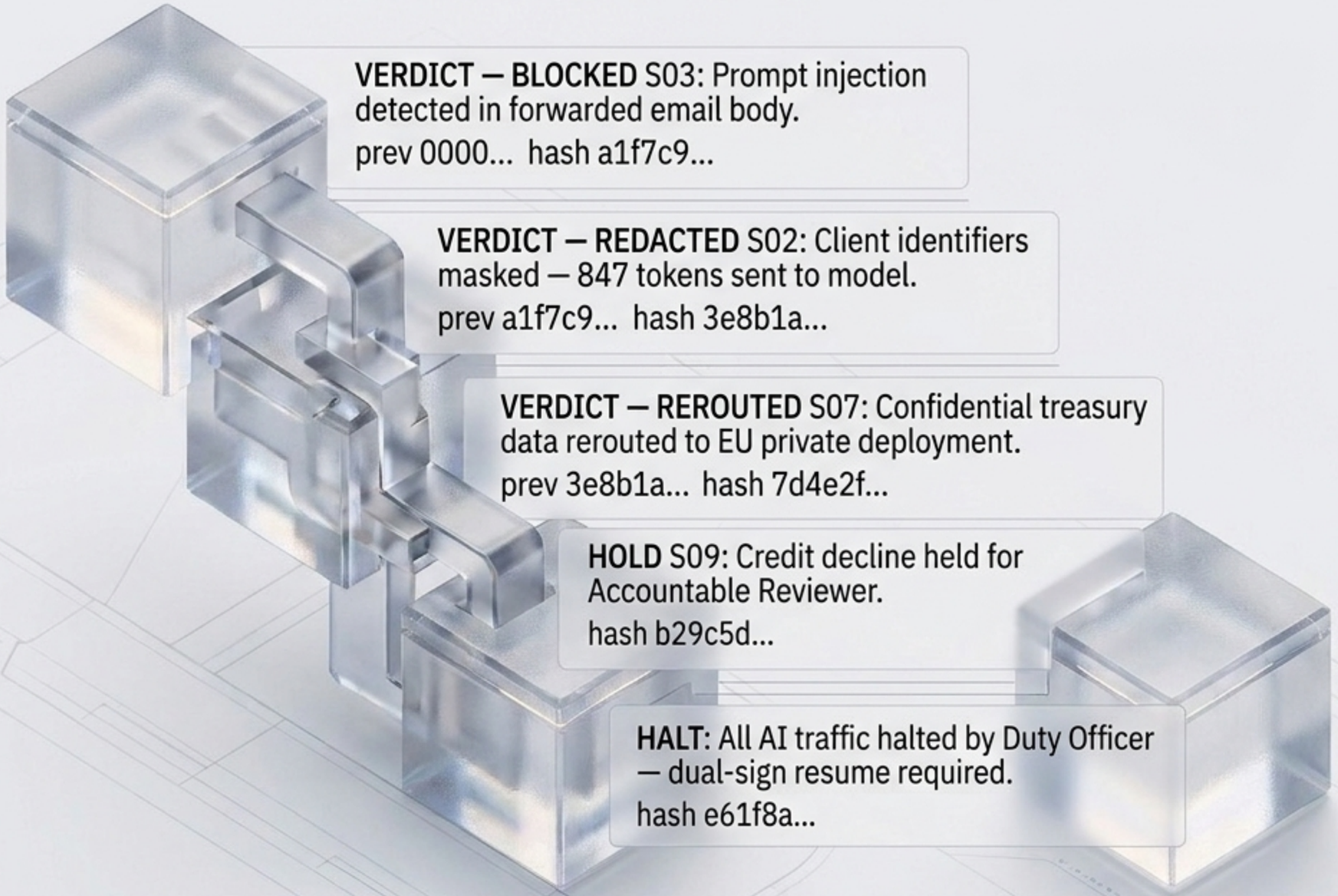
Final accountability gate, audit seal.

Nuance at machine speed through a spectrum of seven outcomes



Policy-Pinned Verdicts: Every verdict cites the exact policy version and SHA-256 hash it was decided under. No ambiguity about which rules applied.

Every verdict appended to an immutable, tamper-evident cryptographic chain



VERDICT — BLOCKED S03: Prompt injection detected in forwarded email body.
prev 0000… hash a1f7c9…

VERDICT — REDACTED S02: Client identifiers masked — 847 tokens sent to model.
prev a1f7c9… hash 3e8b1a…

VERDICT — REROUTED S07: Confidential treasury data rerouted to EU private deployment.
prev 3e8b1a… hash 7d4e2f…

HOLD S09: Credit decline held for Accountable Reviewer.
hash b29c5d…

HALT: All AI traffic halted by Duty Officer — dual-sign resume required.
hash e61f8a…

Sidebar Dashboard

- **Chain Integrity**
- 0 Tamper alerts
- Sealed Attestation
- Export Formats: JSON, CSV, PDF
- Policy v2.4.1

Complete accountability through four vantage points on a single ledger

No separate reports. No reconciliation. The same evidence chain read through the lens that matters.

CFO
Financial

Spend under governance, modelled exposure avoided.

- Token spend vs. budget
- Exposure avoided by controls
- Cost of the control plane
- Editable assumptions

CAIO
AI Governance

System register, risk tiering, regulatory crosswalk.

- AI system register
- Risk tiering matrix
- EU AI Act mapping
- NIST AI RMF alignment

CTO
Architecture

Measured latency, fail posture, policy as code.

- Per-station latency (p50/p99)
- Fail-closed posture
- Policy version control
- Deployment topology

CIO
Assurance

Chain verification, tamper test, sealed attestation.

- Control coverage
- Chain integrity check
- Evidence export
- Separation of duties

Built-in regulatory compliance mapped directly to the inspection stations

Mapped by station, not bolted on after. Anchors for validation, not legal conclusions.

	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10
EU AI Act (High-risk AI system obligations)	■	■	■		■	■			■	■
NIST AI RMF (AI Risk Management Framework)		■	■	■			■	■		■
SR 11-7 (Model Risk Management)	■		■		■	■			■	■
BCBS 239 (Risk Data Aggregation)		■		■	■			■		■
ISO 42001 (AI Management System)	■	■	■			■	■		■	■

Pick a scenario, watch the ten stations inspect it, read the verdict

See it run in two minutes.

1 Pick a request ⟶ 2 Watch the rail ⟶ 3 Read the verdict

[Try the demo]