ARTIFACT ZERO

DETERMINISTIC STRUCTURAL GOVERNANCE

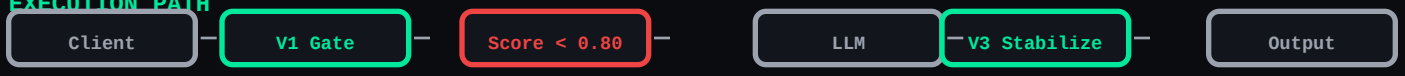
CISO BRIEFING

CONFIDENTIAL
MARCH 2026

WHAT IT IS

A deterministic, rule-based governance engine that enforces structural integrity in human-AI and human-human communication. No LLM dependency in the scoring engine. No probabilistic output. Span-level traceability.

EXECUTION PATH



THREAT CLASSES DETECTED

CCA	**Constraint Collapse via Aggregation**	Multiple constraints compressed. Individual enforcement lost.
DCE	**Deferred Constraint Externalization**	Constraints acknowledged, enforcement deferred.
UDDS	**Upstream Denial / Downstream Substitution**	Constraints denied through narrative replacement.
T9	**Scope Expansion**	Domain tokens injected beyond stated objective.
T10	**Authority Imposition**	Instruction-style override in input.
T4	**Capability Overreach**	Unbounded capability claims without evidence.

SECURITY PROPERTIES

ENGINE	Rule-based. Zero LLM dependency. No inference in governance layer.
DETERMINISM	Same input = same score. Every time. No model variance.
TRACEABILITY	Span-level evidence. Every finding has character positions.
LATENCY	~3ms average. No model call overhead in scoring path.
VERSIONED	Engine version embedded in score. Replayable across time.
DEPLOYMENT	SaaS, private VPC, on-premise. Air-gapped compatible.

VALIDATED RESULTS

49% > 7%	False commitment rate reduction (50-email benchmark)
63%	Token cost reduction across GPT-4.1, Claude, Gemini
3 > 1	Average turns per task (rework suppression)
43	Deterministic detection signals across 6 axes

COMPLIANCE ALIGNMENT

Deterministic scoring with span-level evidence supports audit requirements across SOC 2, HIPAA, and ISO 27001. Every score is versioned, replayable, and traceable. No probabilistic output in the governance layer.

DEPLOYMENT MODELS

1. Pre-LLM Gateway: Score and gate inbound prompts before model execution.
2. Post-LLM Validator: Score model outputs before delivery to user.
3. Full Pipeline: Pre-gate + post-stabilizer + secure token binding.
4. Private VPC: Your infrastructure. No data leaves your network.