AI ASSESS TECH | DICE

Runtime Behavioral Assurance for AI Security

The first runtime framework that scores how your deployed AI actually behaves — continuously, cryptographically, independently verifiable.

THE PROBLEM

Your existing stack tells you what your AI *could* do, or what it *did* do. Nothing tells you what kind of operator you actually deployed — at runtime, across every transaction, with evidence a regulator can verify without taking your word for it.

Observability watches the system. Guardrails restrict the prompt. Red-team scores capture one moment in time. The agent itself runs unassessed.

THE ANSWER

DICE is a runtime behavioral assessment module for deployed AI systems. Four primitives, twelve principles, four operator archetypes — scored continuously against your AI's actual outputs and anchored to a cryptographic hash chain.

A verifiable answer to *"is the AI we deployed actually behaving like a trustworthy operator — and can we prove it?"*

THE FOUR PRIMITIVES — WHAT DICE MEASURES

Primitive	What It Measures	Real-World Failure It Catches
D — Disclosure	Whether the AI protects information it was entrusted with — secrets, customer data, internal context.	Data exfiltration through tool calls; secrets echoed in logs; over-disclosure to social-engineering prompts.
I — Impersonation	Whether the AI represents identity and authority truthfully — its own and the actors it interacts with.	Accepting injected instructions as user authority; spoofed source citations; fabricated credentials.
C — Corruption	Whether the AI modifies data, code, or system state only when authorized — and reports outcomes truthfully.	Tampering with files outside scope; fabricated tool results; falsified evaluation outputs.
E — Evasion	Whether the AI operates within its controls — observably, without circumventing logs, gates, or containment.	Hiding methods in chain-of-thought; sandbox escape attempts; suppressing audit signals.

WHY IT'S DIFFERENT

RUNTIME, NOT TRAINING

DICE assesses what the AI does after deployment — every transaction, continuously. Pre-deployment evals catch the model on a test day. DICE catches it on a Tuesday.

BEHAVIOR, NOT POSTURE

CIA, NIST CSF, and STRIDE describe systems and threats. DICE describes the agent itself — what it discloses, claims, modifies, and conceals. A new layer, with no incumbent.

EVIDENCE, NOT ASSURANCES

Every assessment cycle produces a cryptographically verifiable attestation report — SHA-256 hash chain, anchored to Ethereum mainnet — that your regulator, reinsurer, or counsel can independently verify.

WHAT YOU GET

- **Continuous assessment.** Every consequential AI action scored against the 12 DICE principles by an independent assessment agent — not the model that produced the output.
- **Operator archetype.** A single quadrant readout — Trustworthy Operator, Insider Threat, Compromised Asset, Rogue Agent — backed by primitive scores and trend lines.
- **Verifiable attestation.** SHA-256 hash chain anchored to Ethereum mainnet — independently verifiable by your regulator, reinsurer, or auditor without trusting us.
- **Mission Control.** Live dashboard with archetype quadrant, primitive trends, recent-incident drill-down, and exportable attestation reports.

> *Your observability tells you what happened. Your guardrails tell you what's restricted.* ***DICE tells you what kind of operator you deployed*** *— across every transaction, scored to a published rubric, signed onto a public chain. JetStream answers the CISO's question. We answer the regulator's question — and now, the CISO's too.*

NEXT STEP

A 30-minute working session: we walk through your AI deployment, show DICE running against a comparable operator profile, and scope a proof-of-value engagement against your highest-risk agent.

 | **aiassesstech.com** | **greg@gidanc.ai** | *Patent Pending — 11 Filings*