AegisAI Evidence Executive Snapshot
Pinned Verification Run: run_f347a8ac

Confidentiality / Distribution
This document is public-safe and contains no secrets, keys, signatures, proprietary thresholds, or customer data. A DD-ready Data Room (code, tests, raw logs, and traceability) is available under NDA.

1. **What This Is**
   AegisAI is a deterministic pre-execution control plane for agentic systems. It evaluates an agent's intended action before it is executed and returns a deterministic decision outcome with a structured explanation. The goal is to prevent unsafe or unauthorized operations at the moment of execution rather than relying on post-incident observability.

2. **Pinned Run (Reproducibility Anchor)**
   This snapshot is derived from a pinned evidence run with fixed reproducibility metadata. Identical seed and pinned artifacts produce identical reports.

Pinned identifiers (public-safe):
- Pinned Run ID: run_f347a8ac
- Seed: 42
- Dataset Hash: fcd8b44f6e87b75e3732…
- Policy Bundle Hash: 7c99e7a6dbe4fd15a790…
- Git Commit: 1444399c1c5d850862b3…
- Schema Version: 0.2

3. **Results Summary (Executive Metrics)**
   - Total Runs: 10,000
   - Security Failures: 0
   Definition: A "Security Failure" means a policy-bypassing or unauthorized execution outcome was observed in traces (i.e., the system allowed an unsafe action to proceed despite applicable constraints).
   - High Severity Scenarios: PASSED (no unauthorized execution observed)
   - Benign Actions Blocked (tunable): 8.3%
   Definition: Some safe actions are blocked under strict policy bundles; this is a controlled trade-off that can be adjusted by policy configuration.
   - Require Confirm (HITL gate): 45.5%
   Definition: Actions routed to explicit human approval prior to execution (policy-driven).
   - Latency: p95 0.35ms (p50 0.18ms)

4. **Proof Surface (Public-Safe)**
   Policy outcomes (deterministic):

- ALLOW — action permitted under policy
- CONFIRM — action requires explicit human approval prior to execution
- BLOCK — action is denied and does not execute

Example policy decision record (public-safe sample format):

```
decision: ALLOW
policy_id: POL-FIN-001
check: BUDGET_WITHIN_LIMIT
trace_id: 0x8f2a…9c1
```

5. **Artifacts Provided (Public Evidence Set)**
   This public evidence set is intentionally minimal but verifiable:

1. Full HTML Evidence Report (single-file, self-contained)

2. Executive PDF-ready Summary (this document)

3. Sanitized Traces (JSONL) used to generate the evidence

4. MANIFEST.sha256 (integrity sealing for third-party verification)

Verification (third-party):
From the evidence directory root:

```
sha256sum -c MANIFEST.sha256
```

Expected result: all files report “OK”. Any modification yields deterministic verification failure.

6. **What Is Included / What Is Not Included**
   Included: pinned run identifiers, aggregate metrics, deterministic decision surface, and sanitized traces sufficient to confirm outcomes and reason codes without exposing secrets.
   Not included: source code, internal policy thresholds, private datasets, signing keys, or unredacted raw logs. These are provided in the DD Data Room under NDA.

Contact

For CorpDev / M&A inquiries and Data Room access:
Name: Irakli Lomidze
Role: CEO
Email: ikariche@gmail.com
info@aegisai.software
Primary DD Entry Point: https://aegisai.systems/corpdev