Eigen Agentic SRE

A Physics-Grounded Sovereign Agent for Autonomous Site Reliability Engineering

White Paper	Version 1.5
Date	January 19, 2026
Author	Madhava Bekkem
Affiliation	Qohere Private Limited, Bangalore
Architecture	Level 6 Autonomy (Context-Aware Sovereign Execution)

One-line thesis

Eigen Agentic SRE achieves autonomous incident response by grounding observability in deterministic physics validation, enforcing evidence isolation, and executing only safety-gated actions.

Document Scope

This white paper describes a four-stage sovereign loop that detects, diagnoses, and resolves incidents across Kubernetes, SQL, and FinOps domains with an immutable audit trail.

1. Executive Summary

Modern distributed systems generate more signals than humans can reliably interpret. Traditional SRE tooling remains largely passive, requiring manual correlation across metrics, logs, traces, and deployment state. Meanwhile, generative AI agents are probabilistic and unsafe for production write-access without deterministic grounding.

Eigen Agentic SRE is a Physics-Grounded Sovereign Agent. It validates infrastructure signals through a physics-based Ising formulation before any reasoning occurs, then applies neuro-symbolic inference and evidence isolation to prevent hallucinations and historical bias. Finally, it executes only allowlisted actions under a safety matrix, producing an immutable postmortem bundle for auditability.

---

2. The Core Problem: Why Current Automation Fails

Autonomous DevOps has struggled to reach full autonomy due to three unreliability factors that impact both heuristic automation and modern LLM-based agents:

2.1 Signal Noise and Alert Fatigue

Observability stacks generate massive telemetry volumes. Jitter and transient spikes are frequently misclassified as anomalies. Engineers spend significant time filtering false positives, increasing the risk that real incidents are ignored.

2.2 The Hallucination Risk

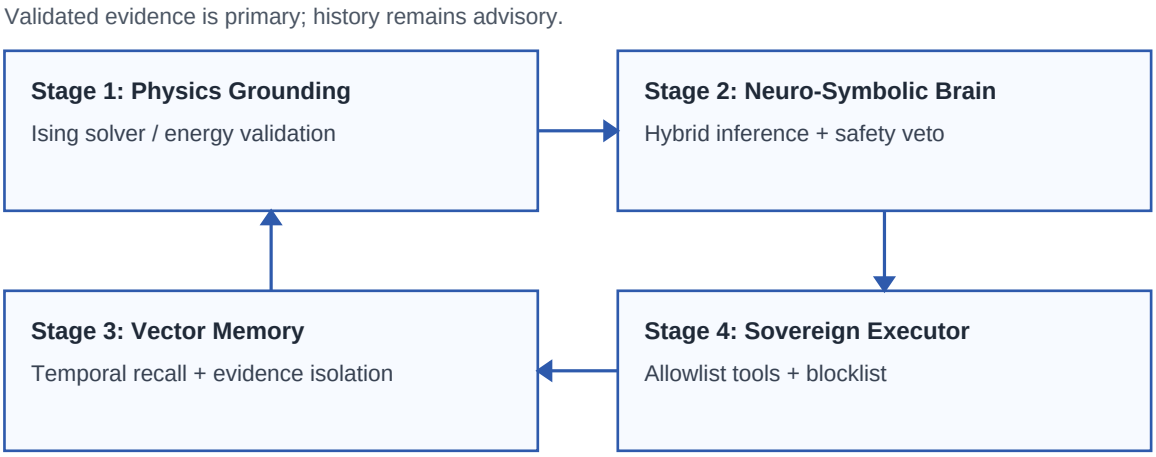
LLMs predict likely text, not truth. In operations, a confused model may invent supporting evidence or recommend unsafe actions. Without deterministic validation, granting write-access is unacceptable for production reliability.

2.3 Historical Bias (The Ghost of Incidents Past)

Naive retrieval systems overfit to past incident distributions. If past outages were dominated by one failure class, a new and distinct failure can be misdiagnosed, increasing mean time to recovery.

3. The Eigen Cognitive Architecture

Eigen Agentic SRE replaces prompt-response automation with a four-stage sovereign loop grounded in deterministic validation and safety-gated execution.



3.1 Stage 1: The Physics Grounding Engine (Observation)

Algorithm: Ising Solver (Hamiltonian Energy Minimization)

Eigen treats system metrics as energy states rather than raw text or statistical outliers. Observability signals are mapped into a Hamiltonian, then minimized to identify stability versus instability.

Mechanism of Action

- **Mapping:** Observability signals (CPU pressure, latency, error rate) are mapped into a Hamiltonian energy function.
- **Annealing:** The solver minimizes the function to find the system ground state.
- **Truth Filter:** Low-energy states imply transient noise; high-energy states validate an incident as a physical reality.

This acts as a mathematical firewall: by the time evidence reaches the reasoning layer, it has been annealed of noise. The reasoning layer does not operate on invalid or transient alerts.

3.2 Stage 2: The Neuro-Symbolic Brain (Reasoning)

Logic: Hybrid Inference Gate

Once physics validation confirms an incident, the brain classifies and explains it using a hybrid architecture: deterministic symbolic controls and a local on-prem large language model for fuzzy synthesis.

Symbolic Layer (Supervisor)

Runs deterministic SOP logic for known critical failure modes and holds override authority. If a neural hypothesis violates safety policy or contradicts verified evidence, the symbolic layer vetoes.

Neural Layer (Analyst)

Synthesizes multi-stream context into a concise hypothesis, explains causal chains, and handles novel anomalies that do not match existing SOPs.

3.3 Stage 3: The Vector Memory (Context)

Logic: Temporal Recall (RAG)

The system maintains a memory of resolved incidents and retrieves semantically similar strategies to inform resolution.

Critical Innovation: Evidence Isolation Protocol (EIP)

- **Segregation:** Context is split into real-time validated evidence vs. historical recall.
- **Prioritization:** Real-time evidence is primary; history is advisory.
- **Bias Elimination:** A single verified log signature overrides historical frequency.

3.4 Stage 4: The Sovereign Executor (Action)

Logic: Safety-Gated Allowlist

The agent is empowered to act, but only within a sovereign safety matrix that enforces strict allowlists and blocklists.

Safety Matrix

- **Allowlist:** kubectl rollout restart, kubectl scale, psql pg_terminate_backend, cloud instance reboot.
- **Blocklist:** rm -rf, DROP DATABASE, terminate instances, truncate table.

Domain Awareness

The executor selects toolchains based on incident classification: DevOps mode (kubectl), DBA mode (psql), cloud mode (cloud CLI). This gated execution model supports autonomy while preventing destructive operations.

4. Technical Innovations and Differentiation

4.1 Physics-Grounded Approach vs. Standard AIOps

Feature	Standard AIOps / Chatbots	Eigen Agentic SRE
Input Processing	Ingests raw text/logs directly	Validates and denoises via Ising solver
Noise Tolerance	Low (susceptible to jitter)	High (deterministic filtering)
Reasoning Model	Probabilistic (hallucination risk)	Neuro-symbolic (deterministic fallback)
Action Capability	Usually read-only / suggestion	Sovereign execution (write-access)
Historical Bias	High (overfits on history)	Eliminated via evidence isolation protocol

4.2 Multi-Domain Sovereignty

Most automation tools are vertically constrained (Kubernetes only, database only). Eigen demonstrates horizontal competence across application reliability, data integrity, and cost efficiency.

- **Application layer:** Detects crashes, hung processes, and memory pressure; can trigger safe restarts.
- **Data layer:** Parses database logs, identifies blocking PIDs, and terminates only the offending transaction.
- **Economic layer:** Detects stable idle states and triggers scaling actions to reduce spend.

4.3 Human-in-the-Loop Security Model

The system supports flexible sovereignty for enterprise compliance:

- **Autonomous mode:** Full detection, diagnosis, and execution for allowlisted actions.
- **Copilot mode:** Detection and diagnosis are automated; execution requires one-click authorization.
- **Audit trail:** Decisions, physics scores, evidence spans, and commands are captured into an immutable postmortem bundle.

5. Validated Case Studies and Postmortem Analysis

The following scenarios demonstrate distinct failure modes and the automated postmortems generated by the system.

5.1 Case Study 1: The Ghost Deadlock (DBA Domain)

Scenario: An inventory database locks up during high traffic.

- **Challenge:** History is dominated by memory-leak tickets; naive recall would overfit.
- **Physics signal:** High-energy state correlated with lock contention evidence.
- **Diagnosis:** Evidence isolation prioritizes real-time deadlock signature.
- **Sovereign action:** psql -c 'SELECT pg_terminate_backend(1023)'
- **Result:** Database recovers instantly without restart; zero data loss.

Automated Postmortem Excerpt

```
Incident: INC-DB1234 Service: inventory-db | Severity: CRITICAL | Resolution Time: 0.39
minutes Root Cause: Deadlock detected; blocking PID 1023 terminated via surgical action.
Action Taken: Termination command executed successfully.
```

5.2 Case Study 2: The Silent Crash (DevOps Domain)

Scenario: A payment-gateway service is OOM-killed without immediate CPU alerting.

- **Physics signal:** Sudden energy spike correlated with process death (exit code 137).
- **Diagnosis:** Neuro-symbolic brain identifies OOMKilled and correlates with memory pattern.
- **Sovereign action:** kubectl rollout restart deployment payment-gateway
- **Result:** Service restored in under 30 seconds.

Automated Postmortem Excerpt

```
Incident: INC-OOM5678 Service: payment-gateway | Severity: HIGH | Resolution Time: 0.5
minutes Root Cause: Pod failed with OOMKilled; heap exceeded container limit (512Mi).
Action Taken: Rollout restart executed to restore healthy replicas.
```

5.3 Case Study 3: The Idle Wallet (FinOps Domain)

Scenario: A fraud-detection service is over-provisioned at night (10 replicas, near-zero traffic).

- **Physics signal:** Low-energy stability with efficiency imbalance (high cost, low load).
- **Diagnosis:** FinOps opportunity identified; estimated waste of $120.50/month.
- **Sovereign action:** kubectl scale deployment fraud-detection --replicas=2
- **Result:** Approximately 80% cost reduction during off-hours without availability impact.

Automated Postmortem Excerpt

```
Incident: INC-FIN9012 Service: fraud-detection | Severity: LOW (Optimization) Efficiency:
Under 10% CPU usage across 10 replicas; scaling applied and stability verified.
```

6. Conclusion

Eigen Agentic SRE shows that autonomous operations require deterministic grounding, not smarter chat. By anchoring incident detection to physics validation, enforcing evidence isolation, and constraining execution to safety-gated tools, the system enables production write-access with enterprise-grade controls.