Synchronicity

Air traffic control for autonomous AI.

A modern airport runs millions of decisions a day — every takeoff, taxi, and landing — without the pilots making those decisions themselves. A separate authority evaluates whether each proposed action is safe, legal, and consistent with the larger system, then signs off before the action happens. Pilots fly the plane. Controllers approve the move.

Autonomous AI needs the same separation. Models decide what to propose. Synchronicity evaluates whether the proposal complies with policy. Execution systems carry out only what was approved. The model doesn't grant its own permissions. The auditor doesn't trust the model's self-report. And every decision leaves a signed, replayable record.

Below are 34 problems the AI governance landscape is trying to solve. One architecture addresses all of them.

Section 1 – Autonomous Agent Behavior

1. An autonomous agent fires a deletion, transfer, or external action it was never authorized to perform

Maps to OWASP AISVS C5.6, NIST AI RMF Govern-3.2, NIST SP 800-207 §2.1 Tenet 6. Every proposed action is canonicalized as a **Structured Proposed Action Descriptor (SPAD)** and submitted to the external governance authority before execution. The authority returns PERMIT, DENY, or HOLD. Execution systems are architecturally prevented from acting without a valid signed decision artifact. Default-deny on failure.

2. An agent invokes tools, APIs, or credentials outside its assigned scope

Maps to OWASP AISVS C9.7, NIST SP 800-207 §3.4.1. SPADs carry tool identity, parameters, and invoking actor. Tool eligibility is evaluated as a deterministic policy rule against the actor's authorized capabilities at evaluation time.

3. An agent chains multiple individually-permitted actions to accomplish a prohibited outcome

Maps to EU AI Act Art. 9 (risk management) and Art. 14 (human oversight), OWASP AISVS C5.6.5. Each action in the chain is independently evaluated. Policy rules can reference prior decision artifacts in the same session, allowing detection of action-sequence patterns that pass individually but collectively violate policy.

4. An agent acts on prompt-injected instructions from untrusted input

Maps to OWASP AISVS C9.7.7, NIST AI 100-2 (Adversarial Machine Learning). SPADs include input lineage metadata. Policy rules can require attested provenance for high-risk actions. Actions originating from untrusted sources are denied at the governance layer regardless of what the model reasoned its way into.

5. An agent escalates privileges beyond its assigned scope

Maps to NIST SP 800-207 §2.1 Tenets 1–3, ISO/IEC 42001 §8.2. Actor scope is bound to policy at evaluation time. The agent cannot self-escalate because the governance authority — not the agent — holds the authoritative actor-to-capability binding.

Section 2 – Decision Provenance and Audit

6. A regulator demands proof of which rules governed a specific AI decision

Maps to EU AI Act Art. 12, NIST AI RMF Measure-2.8, IEEE 7001. Every evaluation produces a signed decision artifact containing the policy version identifier, cryptographic policy content hash, constraint violation identifiers (for DENY outcomes), and a signed timestamp. Reproducible, tamper-evident, deterministically replayable.

7. An auditor needs to reconstruct a historical AI decision exactly

Maps to EU AI Act Art. 12, ISO/IEC 42001 §8.5. A replay verification component re-executes any historical evaluation using the exact policy snapshot active at the original evaluation time. Output matches the original decision artifact.

8. A team needs to detect bias across thousands of AI decisions

Maps to IEEE P7003, EU AI Act Art. 10, NIST AI RMF Measure-2.11. All PERMIT, DENY, and HOLD outcomes are written to a unified append-only log as entries of a common event class. This prevents selective logging and enables population-level bias analysis by downstream analytics. The architecture provides the evidence substrate; bias detection runs on top.

9. Logs are tampered with after the fact to hide a bad decision

Maps to EU AI Act Art. 12, ISO/IEC 27001 A.12.4, NIST SP 800-53 AU-9. Decision artifacts are cryptographically signed at generation. The append-only log architecture prevents retroactive modification. Tampering is detectable through signature verification of any historical artifact.

10. Policy is changed after a decision to make it look compliant in hindsight

Maps to EU AI Act Art. 12, IEEE 7001. Each decision artifact records a cryptographic content hash of the exact policy in force at the moment of evaluation. Subsequent policy changes do not affect the historical record; replay verification will fail if anyone attempts post-hoc substitution.

Section 3 – Multi-Jurisdiction and Multi-Tenant

11. Operating an AI system across EU, US, and sector-specific rules simultaneously

Maps to EU AI Act, NIST AI RMF, sector rules (HIPAA, GLBA, FCRA, FERPA). Policies are composable and version-controlled. Jurisdictional, sector, and organizational policy layers are evaluated with deterministic precedence rules. The decision artifact records which layers contributed to the outcome.

12. Different tenants need different governance rules in the same deployment

Maps to ISO/IEC 42001 §6.1.2, SOC 2 CC6.1. SPADs carry a tenant identifier. Policies are tenant-scoped at evaluation time. Cross-tenant policy bleed is architecturally prevented because evaluation context binds tenant to policy snapshot.

13. Cross-border data flows trigger conflicting governance requirements

Maps to EU AI Act Art. 10, GDPR Art. 44–50, ISO/IEC 27701. Geographic and jurisdictional attributes in the SPAD trigger policy layers specific to data residency, transfer restrictions, and consent state. The decision artifact records which jurisdictional rules were applied.

14. A model is used by a regulated entity but hosted by an unregulated provider

Maps to EU AI Act Art. 28 (deployer/provider obligations), NIST SP 800-161 (supply chain). Synchronicity sits architecturally outside both the model and the host. Governance is enforced at the action layer regardless of who operates the model or where it runs. Compliance posture is portable.

Section 4 – Failure Modes and Resilience

15. The model fails or returns malformed output

Maps to IEEE 7009-2024 (fail-safe design), NIST AI RMF Manage-2.4. Default-deny semantics. If the SPAD cannot be parsed, the policy snapshot is unavailable, or the evaluation engine fails, the

response is DENY or HOLD. Unsafe states are designed to be unreachable under the specified threat model.

16. The governance system itself becomes unavailable

Maps to IEEE 7009-2024, NIST SP 800-53 CP-2. A service availability monitor detects authority unavailability and forces the execution layer into DENY mode. The system fails closed by design rather than degrading to permissive behavior.

17. A policy version cannot be retrieved at evaluation time

Maps to IEEE 7009-2024, EU AI Act Art. 15. A policy snapshot availability monitor detects missing policy versions and returns DENY or HOLD. Actions cannot proceed without a verifiable policy basis.

18. An agent operates in an air-gapped or offline environment

Maps to NIST SP 800-53 SC-36 (distributed processing and storage), IEC 62443 (industrial control systems). Edge-native deployment supports offline decision artifact verification. Governance decisions are evaluable without network connectivity, enabling SCADA, drone, defense, and remote-operation use cases.

Section 5 – Human Oversight

19. Some decisions require a human in the loop, but agents bypass approval

Maps to EU AI Act Art. 14, NIST AI RMF Govern-1.5. Policy rules can require HOLD outcomes for specific action classes. HOLD outcomes route to designated human approvers. Execution cannot proceed without a subsequent PERMIT artifact signed after human approval.

20. A human needs to terminate or override an autonomous process in real time

Maps to EU AI Act Art. 14, ISO/IEC 42001 §6.1.3. Termination authority is a structural element of the architecture. Human operators can revoke policy bindings, force DENY responses, or terminate an entire agent's authorization in real time. Because the governance authority — not the agent — holds the binding, the agent cannot circumvent termination by self-modification.

21. Operators don't know when to intervene because risk signals are buried

Maps to EU AI Act Art. 14, NIST AI RMF Govern-5.1. HOLD outcomes carry structured reason codes for escalation tracking. Operator dashboards subscribe to the unified log and surface escalations in real time with full action context, eliminating the "buried in audit logs" problem.

Section 6 – Performance and Real-Time Constraints

22. Governance evaluation must happen in real time without breaking the user experience

Maps to OWASP AISVS C5.6.5, NIST AI RMF Measure-2.5. The architecture is designed for sub-50 millisecond evaluation latency. Policy resolution is optimized for in-path evaluation, with no I/O dependencies in the hot path.

23. High-throughput AI systems generate too many decisions to govern in line

Maps to ISO/IEC 42001 §8.5. The architecture targets 10,000+ requests per second per node, with horizontal scaling enabled because policy snapshots are the only shared state. Production validation is part of the active development roadmap.

Section 7 – Supply Chain, Provenance, and Identity

24. An AI system uses models or tools from untrusted third parties

Maps to NIST SP 800-161, EU AI Act Art. 25, OWASP AISVS C6. SPADs include cryptographic identifiers for invoked models and tools. Policy rules can require attested provenance. Untrusted artifacts can be denied at the governance layer without modifying the underlying agent.

25. The system needs to prove compliance without disclosing the underlying policy

Maps to EU AI Act Art. 78 (confidentiality) and Recital 71, referencing Directive (EU) 2016/943 on trade secrets; GDPR Art. 5. The decision artifact records what was evaluated and the outcome without requiring exposure of the proprietary policy itself. This protects regulated organizations whose internal governance rules are trade-secret-sensitive.

26. Agent identity is forged or spoofed

Maps to NIST SP 800-63, NIST SP 800-207 §3.3. Every SPAD is cryptographically bound to a specific actor identity at a specific moment in time. Replay attacks fail verification.

27. The underlying model is updated, fine-tuned, or swapped at runtime — policy was written against the old behavior

Maps to NIST AI RMF Manage-2.2 and Manage-4.1, ISO/IEC 42001 §8.4 (AI system impact assessment), EU AI Act Art. 72 (post-market monitoring). Governance operates on the action layer, not the model layer. Two surfaces are observable to policy. First, every SPAD carries the cryptographic identifier of the model that produced it; a change in model identity is visible in the artifact and can trigger re-evaluation, elevated review, or HOLD as a policy decision. Second, behavior-relevant attributes of the proposed action — scope, tools invoked, sensitivity of targets,

output characteristics — are evaluated against policy regardless of which model produced them. A new model that proposes the same out-of-scope action still gets denied. What the architecture does not claim to detect is silent within-version behavioral drift; that is the responsibility of the model-evaluation layer, and Synchronicity is designed to consume signals from it as additional SPAD metadata.

Section 8 – Threats Against the Governance System Itself

28. The governance authority itself is targeted — DoS, policy poisoning, key compromise

Maps to NIST SP 800-53 SI-4, OWASP AISVS C5.6, NIST AI RMF Manage-2.4. The architecture treats the governance authority as a hardened control plane separate from the data plane. Unavailability triggers fail-closed semantics. Policy integrity is cryptographically verified before loading; poisoned policies fail verification and are refused. Key rotation is structural, not optional.

29. Multi-agent orchestration — Agent A delegates to Agent B, which delegates to Agent C

Maps to OWASP Top 10 for Agentic Applications (2025), OWASP State of Agentic AI Security and Governance v2.01 (2026), NIST AI Agent Security RFI 2026-00206. Each agent's actions are independently evaluated against its own policy scope. Delegation does not transfer authority — it creates a new evaluation context bound to the receiving agent's scope. Chains of delegation produce chains of decision artifacts, each independently verifiable.

30. Shadow AI — employees deploy unauthorized models on corporate data without governance

Maps to ISO/IEC 42001 §6.1, NIST AI RMF Govern-1.4. The governance authority can be deployed as a mandatory gateway at the network or identity boundary. Actions originating from unauthorized models or sessions are denied because they cannot produce a valid signed SPAD recognized by the authority. Within the governed network boundary, shadow deployments are visible by their absence of governance artifacts. Bypass via out-of-band channels (personal devices, personal API keys) remains an identity and network-policy problem outside the governance layer's scope.

Section 9 — Standards Conformity and Procurement

31. The organization needs ISO/IEC 42001 certification but lacks technical controls

Maps to ISO/IEC 42001 §6, §8. Architectural components map directly to ISO 42001 controls — SPADs and actor metadata to §8.2 (AI system inventory), the append-only log to §8.5 (monitoring and measurement), policy version resolution to §6.1.2 (risk assessment), fail-closed semantics to §6.1.3 (risk treatment). Synchronicity provides the technical implementation layer that a certifiable management system requires.

32. The organization is bidding on EU AI Act-regulated work and needs a defensible conformity story

Maps to EU AI Act Art. 40 (presumption of conformity via harmonized standards). The architecture is designed to support compliance with EU AI Act Articles 9 (risk management), 10 (data governance), 11 (technical documentation), 12 (logging), 13 (transparency), 14 (human oversight), and 15 (accuracy, robustness, cybersecurity). Final conformity assessment for high-risk systems remains the responsibility of a notified body, but the technical substrate is built to satisfy the underlying requirements.

33. The organization needs NIST AI RMF alignment for U.S. federal procurement

Maps to NIST AI RMF (Govern, Map, Measure, Manage). Policy authoring supports Govern. The SPAD ontology supports Map. The append-only artifact log supports Measure. HOLD and termination authority support Manage. Synchronicity covers the technical control dimensions of the RMF; organizational governance, sociotechnical context, and stakeholder engagement remain the responsibility of the deploying organization.

34. Pre-Action Authorization (PAA) is emerging as a NIST control category

Maps to NIST AI Agent Security RFI 2026-00206. Synchronicity is built around Pre-Action Authorization — the control category NIST is currently scoping in the AI Agent Security RFI. If PAA becomes a formal NIST control, Synchronicity is the architectural reference for what compliance looks like.

How Synchronicity Integrates

The governance authority deploys as a single Go binary. Three integration patterns:

- **Inline gateway** — every agent action passes through the authority before execution. Highest assurance, mandatory for high-risk deployments.

- **Sidecar evaluation** — the authority runs alongside the agent and is consulted by SDK or API. Lower friction, suitable for retrofitting existing systems.
- **Embedded library** — the deterministic engine is linked directly into the host application for offline or edge use cases.

Synchronicity is available for OEM and embedded integration with GRC, AI assurance, and security platforms.

Why One Architecture Addresses All 34

Most governance solutions address one problem at a time — an audit logger here, a tool allowlist there, a content filter somewhere else. Synchronicity is built on the Canonical Responsibility Principle: reasoning systems decide what to propose, the governance authority evaluates whether it complies with policy, and execution systems control how approved actions are carried out. Because every problem above reduces to "did this proposed action satisfy the relevant policy at the moment of proposal," one architecture covers the entire landscape.

The same engine governs a large language model agent, a robotic process automation bot, an algorithmic trading system, or an autonomous agricultural vehicle.

About

Synchronicity is a product of Keystone Digital Holdings LLC. The founder is a contributor to OWASP AISVS (controls C9.7.7, C5.6.5, C5.6.6), has submitted formal responses to the NIST AI Agent Security RFI (2026-00206) and NIST COSAiS, and maintains pending patent applications covering the architecture described above. Patent counsel: Schwegman Lundberg & Woessner.

For an architecture briefing under NDA, contact boone.carlson@keystonedigitalholdingsgroup.com.
