

Nobulex: Case Study for Singapore IMDA Model AI Governance Framework for Agentic AI

Submitted by: Arian Gogani, Maintainer, Nobulex **Contact:** nobulex.dev@gmail.com **Organisation:** Nobulex (nobulex.com) **Date:** April 2026

Overview

Nobulex is an open source cryptographic accountability protocol for AI agents. It addresses a specific gap identified in the MGF for Agentic AI: the need for tamper-evident, independently verifiable records of agent actions, covering both pre-execution authorization and post-execution outcomes.

The Problem the Framework Identifies

The MGF for Agentic AI notes that agents taking autonomous actions create challenges for human accountability, particularly for "unauthorised or erroneous actions" and the difficulty of maintaining "meaningful human control and oversight." Current logging tools produce records that are mutable and self-reported. A log written by the same system that ran the agent cannot prove what the agent actually did.

What Nobulex Implements

Nobulex implements the MGF's "Implement Technical Controls and Processes" dimension through cryptographic receipts. For every agent action:

1. A pre-execution signature commits the agent to a specific action under a specific policy, before execution begins. This proves the action was authorized and defines the scope of what the agent was permitted to do.
2. A post-execution signature binds the actual result to the pre-execution authorization, after execution completes. This proves what the agent did and allows any verifier to confirm the outcome matched the authorization.

Both signatures are hash chained to prior receipts, enabling full session replay and audit by any party holding the public key, without trusting the operator's infrastructure.

Alignment with MGF Dimensions

Assess and Bound Risks Upfront: Nobulex's Covenant primitive allows organisations to define behavioral rules (permit, forbid, require) before deployment. Agents that attempt actions outside their covenant are blocked before execution.

Make Humans Meaningfully Accountable: The bilateral receipt provides an auditable chain of evidence for every decision, enabling post-incident investigation with cryptographic certainty rather than circumstantial logs.

Implement Technical Controls and Processes: Ed25519 digital signatures, SHA-256 hash chaining, and JSON Canonicalization Scheme (RFC 8785) ensure receipts are tamper-evident and independently verifiable.

Enable End-User Responsibility: The open protocol means any organisation can verify receipts independently. There is no dependency on a vendor's verification infrastructure.

Current Adoption

- Merged into Microsoft's Agent Governance Toolkit as the bilateral receipt primitive (PR 1333, April 22, 2026)
- OpenSSF Best Practices Badge passing level (Linux Foundation)
- Receipt format under active review by OpenLineage (Linux Foundation) as an attribution facet
- Referenced in a proposal to the in-toto supply chain attestation framework
- Nineteen independent contributors across LangChain, AutoGen, CrewAI, NousResearch
- Cross-language interoperability verified: TypeScript and Python implementations produce byte-identical digests
- Scheduled for April 30, 2026 cross-implementation test alongside APS, AgentID, AgentNexus, and Concordia

Relevance to Singapore's MGF

The MGF identifies "implementing technical controls throughout the agent lifecycle" as a core dimension. Nobulex provides exactly that: a protocol-level control that operates at the action boundary, before and after each agent decision, producing an independently auditable record. The protocol is MIT licensed and free for any organisation to implement.

Contact

Arian Gogani nobulex.dev@gmail.com github.com/arian-gogani/nobulex nobulex.com