

# Technical Specification: Unified Attribution & Policy Engine

**Version:** 1.0 (Alpha) | **Core Focus:** Deterministic AI Governance

## 1. Objective

To provide a model-agnostic, zero-trust middleware that ensures AI interactions are deterministically anchored to an authorized source corpus. This architecture bridges Enterprise compliance and Civilizational source integrity.

## 2. Core Infrastructure Components

Enterprise: Solving for regulatory auditability and provenance in finance/legal.

Civilization: Establishing the interfaith digital standard for how AI interacts with foundational texts.

Component	Technical Mechanism	Operational Value
Merkle-Anchored Integrity	Content-addressable hashing of source texts to eliminate "textual drift."	Prevents hallucinations or unauthorized edits to the core text.
Policy Maniest (JSON)	Deterministic logic injection defining transformation rules and metadata.	Allows communities/firms to set hard guardrails on AI behavior.
Semantic Guardrails	Post-generation vector distance analysis to verify source alignment.	

## 3. The "Sacred Trace" (Auditability)

Every response generates a unique Traceability Hash:  
Hash(Query + Policy\_ID + Source\_Root + LLM\_Output)

This creates a cryptographic receipt allowing any entity (auditor, leader, or user) to verify the chain of custody from the source text to the AI output. This makes the logic "replayable" and transparent.

"Governance infrastructure not interpretive authority."