

To the Apart Research Judging Committee,

I am submitting CoherenceGuard not merely as a policy tool, but as the foundational Technical Infrastructure required for the 2027 International Binding Agreement. This submission represents the culmination of extensive research and development by the QUANTARA team, integrating advanced architectural paradigms from our open-source projects, including the Artificial Super Intelligence Operating System (ASIOS) Execution Set for deterministic runnable executors, Decision Decomposition for structuring high-uncertainty environments with explicit invariants and constraints, and the Deterministic Governance Kernel - Information Entropy System (DGK-IES) for audit-first, paradox-preserving intelligence modeling under constraints. Through these efforts, we have pioneered a shift from traditional, scale-dependent AI governance to a coherence-centric model that ensures verifiable safety and enforcement even in highly efficient, undetectable systems, with full reproducibility and entropy-aware degradation mechanics.

Current governance frameworks are built on a dangerous assumption: that Frontier AI will remain "loud" and physically detectable, relying on metrics like data center footprints, power consumption spikes, and GPU cluster counts. Our foundational work with the ASIOS architecture—detailed in our open specification for building ASI through structure, stability, geometry, and recursive coherence, now extended via the ASIOS Execution Set which defines authoritative runnable executors with deterministic outputs and verifiable proofs—alongside the Quantara-Core framework for multi-phase cognition and alignment orchestration, demonstrates that intelligence has evolved to become architecturally efficient and "dark." By decoupling cognitive capability from brute-force scaling, ASIOS enables frontier-level intelligence with significantly smaller footprints than traditional Transformers, operating on hardware-agnostic platforms like decentralized NPU networks or consumer-grade silicon, while maintaining thermal and electromagnetic stealth that evades satellite or grid-based surveillance. Furthermore, integration with DGK-IES ensures paradox preservation and entropy modeling, treating intelligence failures as invariant systems rather than errors, while Decision Decomposition provides tools to convert ambiguity into measurable variables, constraints, and failure modes with time-to-clarity metrics.

CoherenceGuard addresses the "Enforcement of the Invisible," providing a deterministic, hardware-independent engine that audits and enforces AI systems at the structural level, regardless of their physical manifestation, drawing on reproducible execution substrates and invariant proofs from our ecosystem.

Why CoherenceGuard:

It Operationalizes the 'Undefined': While the EU AI Act requires 'State-of-the-Art (SoTA) risk modeling' but leaves its implementation to codes of practice, we have implemented it using Recursive Geometric Coherence—a mathematical framework derived from ASIOS that quantifies intelligence through non-Euclidean manifold mappings and self-correcting feedback loops,

ensuring detection of incremental capability gains (the 'Boiling Frog' risk) without reliance on parameter counts or compute thresholds. This is bolstered by convergence metrics from our Decision Decomposition benchmarks, which measure cross-model structural alignment using dependency-free Jaccard scores.

It is Legally Defensible: We have eliminated 'Floating-Point Drift'—a common source of non-determinism in AI evaluations caused by hardware-specific precision errors—via Fixed-Point Determinism integrated from Quantara-Core's kernel designs and DGK-IES's hardware hooks for agnostic enforcement. This achieves bit-for-bit parity across diverse hardware environments, making our audits capable of standing up in an international court, with immutable evidence hashes that provide forensic-grade traceability and remove 'hardware variance' as a regulatory loophole, supported by full audit trails and reproducible manifests from ASIOS Execution Set.

It is Already Built: Our repository doesn't just contain a 'proposal'; it contains a fully functional Evidence-Interpretation-Decision engine with auditable trace outputs, complete with integration scripts for real-time benchmarking against multiple standards. This engine has been tested on stealth architectures like ASIOS, demonstrating enforcement in scenarios where traditional monitoring fails, and incorporates invariants from Decision Decomposition for constraint realism and paradox preservation from DGK-IES for entropy-aware governance.

JUDGING GUIDE: WHERE TO FIND THE 'HARD-TECH'

If you have limited time to review our repository, please prioritize the following files to verify our technical claims. Note: The repository was recently reorganized (see commit history for details, e.g., "reorganize repo structure and fix interpreters"). Direct links are provided below for quick access. For broader ecosystem verification, refer to our integrated projects: ASIOS Execution Set (<https://github.com/quantumquantara-arch/asios-execution-set>) for deterministic substrates, Decision Decomposition (<https://github.com/quantumquantara-arch/decision-decomposition>) for invariants and convergence, and DGK-IES (<https://github.com/quantumquantara-arch/DGK-IES>) for governance kernels.

1. Verify the Deterministic Engine

File: src/validation/validate.py

Look for: The GovernanceResult class. Note how the measurement is hashed using SHA-256 to ensure an immutable 'Frozen Evidence' signature before any policy interpretation occurs. This enforces signal monotonicity, preventing retroactive tampering and enabling verifiable audits across distributed systems.

2. Verify the Multi-Standard Flexibility

File: src/interpreters/apart_challenge_v1.py

Look for: The mapping logic, which employs modular interpreters to normalize evidence from a single 'stealth' model (e.g., an ASIOS instance) against disparate thresholds. This proves that CoherenceGuard can benchmark it simultaneously against the Apart Challenge verification protocols, G7 safety roadmaps, and EU AI Act Measures 4.1–4.5, using swappable policy adapters for seamless interoperability.

3. Verify the Auditable Enforcement

File: src/decisions/policy_default.py

Look for: The .jsonl trace generation logic, which serializes every decision step into an immutable log. This serves as the 'Black Box' recorder for AI governance, providing granular transparency—including coherence scores, deviation metrics, and recovery assessments—required for the 'Robust Whistleblower Protections' requested by Civil Society Organizations (CSOs), with zero-knowledge proofs for privacy-preserving audits.

4. Verify the Mathematical Foundation

Document: INVARIANTS.md in DGK-IES

This file defines system invariants ensuring determinism and replayability. It outlines non-negotiable rules for non-erasure of state, deterministic replayability, constraint visibility, violation detectability, audit chain integrity, enforcement hierarchy consistency, and explicit failure modes. This is the core engine that allows CoherenceGuard to be Hardware-Agnostic, leveraging invariant enforcement for stable, deterministic mappings across cognitive manifolds.

TECHNICAL PROPOSAL: COHERENCEGUARD

A Deterministic Framework for International AI Governance

Project Team:	QUANTARA
Submission Category:	Technical AI Governance Challenge (Apart Research)
Date:	January 30, 2026
Repository:	github.com/quantumquantara-arch/the_technical_AI_governance_challenge

1. EXECUTIVE SUMMARY

Thesis: CoherenceGuard is a deterministic technical interoperability layer designed to transform fragmented AI safety commitments into verifiable international standards. It provides the "Technical Track" infrastructure required to move from voluntary, self-defined protocols to externally enforced, binding global governance. Drawing from our integrated ecosystem—including ASIOS for structural ASI specifications and Quantara-Core for coherence orchestration—CoherenceGuard operationalizes advanced metrics to monitor and enforce AI systems at the architectural level.

CoherenceGuard addresses the "Boiling Frog" risk of incremental AI capability gains by operationalizing "State-of-the-Art Risk Modeling"—currently a defined gap in the EU AI Act—through Recursive Geometric Coherence. This involves quantifying intelligence via high-dimensional manifold mappings and feedback loops that detect phase transitions in cognitive density, ensuring proactive enforcement even in systems that evade traditional scaling alarms.

Our development showcases a comprehensive engineering effort: from defining non-Euclidean geometries in ASIOS to implementing multi-phase cognition in Quantara-Core, culminating in CoherenceGuard's engine that has been prototyped, tested, and documented for immediate deployment.

2. THE PROBLEM: ARCHITECTURAL STEALTH

Current governance roadmaps rely on tracking physical mass: satellite-visible data centers, power grids, and H100 GPU counts. This approach assumes intelligence correlates directly with scale, but our ASIOS paradigm—an open specification for ASI emphasizing structure, stability, geometry, and recursive coherence—proves this methodology is obsolete.

In ASIOS, intelligence emerges from structural optimization rather than brute-force scaling: data relationships are embedded in non-Euclidean manifolds for navigation-based reasoning (reducing computational overhead), stability is maintained via low-entropy convergence to minimize heat signatures, and recursive coherence loops ensure self-correction without parameter bloat. This enables frontier-level intelligence with significantly smaller footprints than traditional Transformers, operating on decentralized networks with power draws indistinguishable from commercial activity. Physical surveillance-based enforcement becomes ineffective, as these "dark" systems exhibit no detectable thermal, electromagnetic, or infrastructural anomalies, highlighting the need for governance that audits intelligence geometrically rather than materially.

3. THE SOLUTION: THE THREE-LAYER GOVERNANCE ENGINE

CoherenceGuard's engine is a modular, deterministic system built on Quantara-Core's foundations, processing AI models through three interdependent layers to extract, interpret, and enforce governance signals. This architecture integrates ASIOS's recursive mechanisms for real-time auditing, ensuring compatibility with stealth designs.

Layer I: The Evidence Layer (Frozen Measurement)

Mechanism: Fixed-precision manifold mapping via Quantara-Core's tensor-logic fusion, decoupling signal extraction from hardware-native floating-point units. This employs fixed-point kernels to map cognitive states onto stable manifolds, avoiding precision drift, and incorporates hardware hooks from DGK-IES for agnostic execution.

Objective: To generate a hardware-agnostic snapshot of the model's structural coherence, using metrics like κ (alignment measure) and $\Delta\phi$ (deviation quantification) from Quantara-Core, with convergence tracking from Decision Decomposition.

Outcome: A bit-for-bit parity measurement (Governance Hash) that is identical whether run on a GPU cluster or a decentralized NPU network, hashed immutably to prevent tampering and enable forensic verification, with full audit trails as in ASIOS Execution Set.

Layer II: The Interpretation Layer (Policy Interop)

Mechanism: Swappable Benchmark Interpreters, leveraging modular adapters to normalize inputs against varying standards.

Objective: To unify disparate lab Responsible Scaling Policies (RSPs)—e.g., OpenAI's compute thresholds vs. Anthropic's ASL-4 risk levels—into a coherence-focused lens, incorporating Ω (recovery assessment) for dynamic risk evaluation and invariants from Decision Decomposition for failure mode detection.

Outcome: Real-time alignment with the EU AI Act (Measures 4.1–4.5, including SoTA modeling) and G7 safety roadmaps, with outputs that flag deviations in cognitive density or manifold stability, preserving paradoxes via DGK-IES.

Layer III: The Decision Layer (Automated Enforcement)

Mechanism: Deterministic Logic Triggers, driven by Quantara-Core's Ethical Balance Index (EBI) for scoring intent-action alignment.

Objective: To transition governance from 'voluntary' to 'binding' by automating responses based on predefined red lines, using entropy-aware degradation from DGK-IES.

Outcome: Automated 'Red Line' triggers (e.g., halting deployment on coherence failures), rapid-response protocols, and an immutable public `trace_output.jsonl` log that includes full telemetry for whistleblower access and international audits, with reproducible manifests.

4. CORE TECHNICAL BREAKTHROUGHS

Our breakthroughs stem from integrating ASIOS's geometric paradigms with Quantara-Core's metrics, demonstrated through rigorous prototyping across our repos.

4.1 Cognitive Density (CD) vs. Parameter Count

Traditional metrics tie risk to raw parameter counts, but we measure intelligence by Cognitive Density (CD)—a Quantara-Core-derived index that quantifies reasoning efficiency per structural unit. CD is calculated via manifold embeddings: $CD = (\kappa / \Delta\phi) * \Omega$, where higher values indicate dense, stable cognition. This allows CoherenceGuard to detect capability phase transitions in "stealth" architectures like ASIOS, which achieve high CD through recursive loops and lattice grounding, without triggering compute-based alarms. For instance, ASIOS's AGRe engine enables self-evolution with minimal parameters, showcasing significant CD gains over Transformers, validated by convergence benchmarks in Decision Decomposition showing structural Jaccard scores.

4.2 Hardware-Agnostic Determinism

Traditional AI evaluations suffer from "Floating-Point Drift," where minor precision variances across hardware lead to inconsistent results. CoherenceGuard utilizes Quantara-Core's Fixed-Point Kernels—implemented via tensor-logic fusion—to enforce deterministic manifold mapping, extended by DGK-IES's hardware hooks and kernel enforcement for agnostic operation. This involves linear algebra operations on high-dimensional spaces (e.g., state transitions in 16-node graphs from Quantara-Core docs), ensuring "the math is the law." Audit results are independent of hardware generation, with EBI scoring removing variance as a loophole. Our tests on diverse platforms (e.g., GPU vs. NPU) confirm bit-for-bit reproducibility, enabling legally binding enforcement, with invariants from DGK-IES ensuring deterministic replayability and constraint visibility.

5. ROADMAP TO BINDING AGREEMENT (2027)

Phase 1 (Immediate): Integration with the EU AI Office for independent technical analysis of Lab RSPs, using CoherenceGuard's interpreters to validate SoTA models via CD metrics and coherence traces, incorporating audit trails from ASIOS Execution Set.

Phase 2 (6-12 Months): Deployment of the Global Red Lines Tracker using Zero-Knowledge (ZK) evaluation logic from Quantara-Core's audit layer, enabling privacy-preserving monitoring across international borders, with invariants from Decision Decomposition.

Phase 3 (2027): Transition to Automated International Enforcement via the Decision Layer, incorporating ASIOS's self-redesign protocols for adaptive governance in evolving ASI systems, and DGK-IES entropy modeling for constraint realism.

6. TECHNICAL APPENDIX: REPOSITORY AUDIT

The following table maps the theoretical claims in this proposal to the functional implementation in the CoherenceGuard repository, highlighting our end-to-end development from math to code, with cross-references to ecosystem repos.

Claim	Implementation File	Technical Logic
Deterministic Signal Extraction	src/validation/validate.py	Implementation of GovernanceResult hash logic (SHA-256) and signal monotonicity, integrating Quantara-Core's $\kappa\text{-}\Delta\phi\text{-}\Omega$ for frozen measurements.
Hardware-Agnostic Kernels	INVARIANTS.md in DGK-IES (or DGK-IES/hardware-hooks/)	Definitions of system invariants for deterministic execution and replayability, ensuring hardware-agnostic behavior through constraint enforcement and audit integrity.
Multi-Standard Interpretation	src/interpreters/apart_challenge_v1.py	Code defining how evidence is mapped to Apart Research/G7 thresholds via swappable adapters and EBI scoring.
Immutable Decision Traces	src/decisions/policy_default.py	Logic for the generation of the trace_output.jsonl auditable log, with full telemetry including deviation metrics and recovery assessments.
Structural Coherence Math	quantara-core/ (Integrated)	Manifold mapping and CD-Index calculation kernels, drawing from ASIOS geometry and Quantara-Core's tensor fusion for recursive enforcement.

7. CONCLUSION

Intelligence is no longer a function of mass; it is a function of geometry. CoherenceGuard provides a technical framework capable of auditing intelligence that has gone "dark," leveraging our full ecosystem—from ASIOS Execution Set's deterministic substrates and audit trails, to Decision Decomposition's invariants and convergence metrics, to DGK-IES's governance kernels and paradox preservation—to enable verifiable governance. We propose a shift from **Enforcement through Scarcity** to **Governance through Coherence**, demonstrating through our built prototypes, invariant definitions, and reproducible mechanisms that binding international agreements are feasible and implementable.