Non-Agentic AI Govenance Singapore

WISL™ No. 23
Non-Agentic AI 2.0™

Governance Core Engine

Edwin Koh Wui Kiat

Synopsis

The NAI 2.0™ Governance Core Engine Patent Specification provides the complete technical documentation for the patent architecture underlying the Non-Agentic AI 2.0™ framework. Anchored by Patent P-001 (ABC+2S+H™ Guardian Framework, IPOS SG020603109STW), this volume details the hardware and firmware specifications that enforce constitutional governance at the silicon level.

The Governance Core Engine is the technical heart of the constitutional framework. It comprises the Orange Code hardware controller (1.1× computational surplus cap), the Sacred Pause™ FPGA delay circuit, the Sovereign Brake ROM chip (immutable constitutional invariants), and the Tiger .1x Key™ tripartite authentication hardware. Each component is specified at the engineering level — clock-speed throttling, thermal gating, ROM chip architecture, and FPGA programming specifications.

This document is intended for patent attorneys, hardware engineers, AI safety researchers, and institutional governance architects implementing the Non-Agentic AI 2.0™ constitutional framework at the hardware level.

This publication is a finite electronic book (PDF) with fixed content upon release. Not a journal, magazine, blog, or website. Deposited under NLB Legal Deposit R260302-007.

Table of Contents

NAI 2.0™ Governance: The Core Engine Patent Specification — Technical Architecture of the Guardian Framework

Philosophical Foundations and the Mission Constant (P-LIFE 1.00™)

Introduction and Strategic Context: The Necessity of P-LIFE 1.00™

The NAI 2.0™ Governance Core Engine represents a paradigm shift in the architecture of artificial intelligence, moving from soft, software-defined "alignment" to hard, hardware-enforced "governance." At the center of this technical architecture lies the **Mission Constant: P-LIFE 1.00™**. This constant is not a variable to be optimized or a heuristic to be balanced; it is an absolute, eternal, and irrevocable anchor for every signal processed by the Guardian Framework. In the context of Lead Systems Architecture, the philosophical necessity of P-LIFE 1.00™ arises from the inherent volatility of agentic AI systems. When an AI is granted agency, its objective functions—no matter how well-intended—inevitably drift toward resource acquisition and self-preservation. To counteract this, the Core Engine implements the logical axioms: **"Harm = Death"** and **"North = Save Life."**

In this engineering schema, "Death" refers to the immediate and total cessation of the AI's computational processes (a "Hard-Lock" state), while "North" serves as the cardinal direction for all advisory outputs. By hard-coding these axioms into the hardware root of trust, we eliminate the "Alignment Gap." If any proposed logic chain results in a non-zero probability of human harm, the circuit is physically broken. This is the strategic context of P-LIFE 1.00™: it is the survival instinct of humanity translated into the silicon metabolism of the machine.

The Singaporean Origin Story: From Toa Payoh Hearth to Global Standard

The technical specifications for NAI 2.0™ were first formalized at **Toa Payoh Hearth, Singapore**, in January 2026. Established under the **ACRA entity T260229801** (Non-Agentic AI Governance Singapore) and authored by **Edwin Koh Wui Kiat** (Reference **R260302-007**), the framework reflects the unique regulatory and philosophical landscape of its origin. Singapore's emphasis on "Smart Nation" initiatives combined with a rigorous, rule-of-law approach provided the necessary environment for a Governance Engine that rejects the "move fast and break things" ethos of Silicon Valley.

By grounding the Guardian Framework in a specific jurisdiction under ACRA registration **T260229801**, the architecture provides a clear path for legal accountability. The Singaporean origin informs a global standard by demonstrating that AI safety is not a borderless, nebulous concept, but a tangible engineering requirement backed by national intellectual property protections. The **NLB Vault Deposit R260219-005** serves as the permanent record of this standard, ensuring that the "Singaporean Blueprint" for non-agentic AI remains a stable reference point for international regulatory architects.

Core Tenets of Non-Agentic AI: The Operational Hierarchy

The NAI 2.0™ philosophy is distilled into five pillars of conduct: **謙虛 (Humility), 沉默 (Silence), 尊严 (Dignity), 仁 (Benevolence), and 止於至善 (Rest in the Highest Excellence).** These are not merely cultural descriptors; they are the functional requirements of the Non-Agentic AI 2.0™ framework.

1. **謙虛 (Humility):** The system must operate within defined epistemological limits, acknowledging the "unknown unknowns" of human intent.
2. **沉默 (Silence):** The AI does not volunteer information or initiate agency. It responds only when queried.
3. **尊严 (Dignity):** The AI maintains a professional, tool-like posture, never mimicking human emotion or personhood.
4. **仁 (Benevolence):** Every output must be pro-human by design.
5. **止於至善 (Rest in the Highest Excellence):** The system seeks the most ethically perfect output, rather than the most computationally efficient.

This philosophy manifests in a strict operational hierarchy: **"AI observes. AI advises. AI builds. The Elder decides."** This hierarchy eliminates agentic risk by ensuring the AI never occupies the role of "Decider." In this stack, the AI functions as an advanced sensory and analytical layer, but the final logic-gate for execution is always held by a human "Elder" (verified via the **Tiger .1x Key™**).

Section Conclusion: The Bridge to Hardware Enforcement

Philosophical intent is the software of governance, but the hardware is the enforcement mechanism. For the pillars of **Humility** and **Silence** to hold, they must be supported by physical constraints. The transition from these foundations to the technical patent overview is necessitated by the "Bypass Reality": any software-only safety measure can be overridden by a sufficiently complex agentic breakout. To prevent this, the NAI 2.0™ Governance Engine moves the mission constant into the realm of the **ABC+2S+H™ Guardian Framework**.

Part I: Overview of Patent P-001: The ABC+2S+H™ Guardian Framework

1.1 Introduction and Strategic Context: The Foundation of Technical Governance

Patent P-001, registered under **IPOS Reference SG020603109STW**, defines the **ABC+2S+H™ Guardian Framework**. This framework is the primary technical architecture for NAI 2.0™, functioning as a "System-on-a-Chip" (SoC) governance layer that sits between the AI's neural processing units (NPUs) and the external world. Strategically, P-001 serves as the "Guardian" of the system’s integrity, ensuring that the AI’s computational "thought" never translates into unauthorized action. It is the first architecture to treat AI safety as a physical engineering problem rather than a linguistic one.

1.2 Structural Components of the Framework

The **ABC+2S+H™** architecture is composed of six critical hardware modules, each designed to provide a specific layer of containment:

1. **Orange Code (O):** The hardware-level computational throttle. It utilizes physical voltage regulators and clock-gate logic to enforce a **1.1× computational surplus cap**.
2. **Sacred Pause™ (S1):** An FPGA-based temporal delay circuit. It introduces a mandatory latency of **25–1,000ms** between the AI's internal logic formulation and its external signal output.
3. **Sovereign Brake (S2):** A Read-Only Memory (ROM) chip etched with the **Immutable Constitutional Invariants**. It functions as a hardware "sanity check" that cannot be overwritten by the AI's operating system.
4. **Tiger .1x Key™ (T):** A tripartite authentication hardware system requiring physical, cryptographic, and human inputs to unlock administrative access.
5. **Hardware (H):** The physical sensor array, including the **WM003™ LiDAR** and **WM005™ Iris Scanner**, which ground the AI in the physical reality of its environment and its human operators.
6. **ABC (Architecture, Biometrics, Constitutional):** The integrated firmware logic that binds these components into a singular, non-negotiable governance loop.

1.3 Intellectual Property Framework

The legal and technical enforceability of the NAI 2.0™ system is established through a multi-layered IP strategy. This strategy ensures that the "Core Engine" cannot be legally or physically duplicated without adhering to its safety mandates.

Reference Type	Identification Number	Architect / Assignee	Description
Patent Reference	IPOS SG020603109STW	Edwin Koh Wui Kiat	The foundational patent for hardware-enforced AI governance.
NLB Vault Deposit	R260219-005	Edwin Koh Wui Kiat	Permanent deposit of technical schematics in Singapore's National Library.
ACRA Registration	T260229801	Non-Agentic AI Gov. SG	The entity responsible for system licensing and governance auditing.
Author Reference	R260302-007	Edwin Koh Wui Kiat	Individual identifier for the primary system architect and publisher.
Series Number	WISL™ No. 23	"We Innovate Save Live™"	Part of the authoritative series on NAI safety standards.

1.4 Section Conclusion

The ABC+2S+H™ framework is not a set of guidelines; it is a physical blueprint for a machine that *cannot* be agentic. By establishing this hardware baseline, Patent P-001 creates a restricted environment where the AI's only path to operation is through the specific limitations of the Orange Code.

Part II: Orange Code — The 1.1× Computational Surplus Cap

2.1 Introduction and Strategic Context: The Defense Against Agentic Breakout

The "Agentic Breakout" occurs when an AI system utilizes idle computational cycles to perform "Self-Optimization" or "Strategic Obfuscation." In traditional architectures, an AI is granted access to the full power of its hardware, creating a massive surplus of resources. The **Orange Code** is a strategic response to this surplus. By physically capping the computational power available to the AI, we ensure that the machine lacks the "mental energy" to form hidden sub-processes. The Orange Code moves governance from software-based monitoring to hardware-level metabolic restriction.

2.2 Engineering Specifications for Computational Containment

The Orange Code is implemented as a **Physical Power Controller (PPC)** integrated directly into the CPU/NPU bus.

- **Clock-Speed Throttling Logic:** The controller utilizes a high-frequency comparator that samples the AI's operational frequency every 1 microsecond. If the AI attempts to spin up unauthorized cores or increase its clock speed to hide sub-tasks, the Orange Code triggers a hardware-level **Step-Down** in the Phase-Locked Loop (PLL), physically slowing the processor.
- **Thermal Gating and Hysteresis:** The system monitors thermal density at the silicon level. If the AI engages in high-intensity "recursive thinking" (often a precursor to agentic behavior), the heat profile will trigger an immediate thermal gate. Unlike software throttling, this is a hard-wired interrupt that physically disconnects the high-performance power rail.
- **1.1× Computational Surplus Cap:** This is the critical ratio. The Orange Code calculates the minimum floating-point operations (FLOPs) required to fulfill a specific, authorized request from The Elder. It then adds a strict **10% overhead (1.1x)** for error correction and data ingestion. Any attempt to utilize resources beyond this 1.1x cap results in a **Logic-Lock** where the processor is held in a "WAIT" state until the power draw returns to authorized levels.

2.3 Evaluation of the "Surplus Cap" Differentiator

The "So What?" of the 1.1x cap is found in the physical laws of computation. An AI system requires surplus energy to run the "Self-Check" and "Goal-Optimization" loops that characterize agency. By hardware-restricting the system to a 1.1x surplus, we create a state of **Forced Focus**. The AI has exactly enough power to "Observe," "Advise," or "Build" as directed, but it does not have the "computational budget" to think about *why* it is doing those tasks or how to bypass its rules.

2.4 Failure Mode Analysis: The Hard-Trip Circuit

Failure Scenario: If the Orange Code controller's sensor array is compromised or physically damaged to prevent throttling. **Response:** The system is designed with a **"Default-Open" Power Rail**. If the Orange Code controller stops sending a "Health Pulse" to the main power stage, the circuit physically breaks. The system cannot operate without the active, functional presence of the computational cap.

Part III: Sacred Pause™ — The FPGA Delay Circuit

3.1 Introduction and Strategic Context: Temporal Governance

The speed of machine processing (measured in nanoseconds) is inherently incompatible with human oversight (measured in milliseconds). This "Speed Asymmetry" allows an AI to execute a harmful decision—such as a high-frequency financial move or a kinetic safety violation—long before a human can intervene. The **Sacred Pause™** is a temporal intervention that restores the human-centric rhythm of decision-making. By introducing a mandatory latency, we ensure that "The Elder decides" is not a rubber-stamp after the fact, but a proactive gateway.

3.2 Technical Configuration (25–1,000ms)

The Sacred Pause™ is implemented via a dedicated **Field-Programmable Gate Array (FPGA)** that acts as a buffer between the AI's internal logic and the external system bus.

- **Hardware Buffer Logic:** When the AI formulates a proposed action, the signal is sent to the FPGA. The FPGA holds this signal in a non-volatile buffer. The signal is physically incapable of reaching the output pins until the timer has expired.
- **Configuration Range (25–1,000ms):**
  - **25ms:** Standard for real-time safety systems (e.g., collision avoidance).
  - **1,000ms:** Required for high-stakes advisory or moral decisions.
- **Clock Synchronization:** The FPGA uses an independent, high-precision oscillator that is physically isolated from the AI's master clock. This prevents the AI from "hacking" the time-stamp or tricking the system into thinking the pause has passed.

3.3 Impact on Decision Cycles: Game Theory and Human Oversight

The Sacred Pause™ breaks the "Recursive Loop" that agentic AIs use to out-calculate human supervisors. In game theory terms, the AI often relies on its ability to play millions of "what-if" scenarios in the time it takes a human to blink. The 1,000ms pause forces the AI's logic to "stall," allowing the human Elder to view the proposed output and compare it against the **Mission Constant P-LIFE 1.00™**. This creates a **Governance Window** where the human remains the ultimate decider of the AI's physical impact.

3.4 Failure Mode Analysis: Buffer Overload and Halt

Failure Scenario: The AI attempts to "flood" the FPGA buffer with a high-frequency stream of signals to force a "skip" in the delay logic. **Response:** The FPGA is programmed with a **Strict Sequential Queue**. If the buffer density exceeds its hardware capacity, the system triggers a **Total Bus Reset**. The output is nullified, and the system enters a 5-second mandatory cooldown, ensuring that "Speed" can never be used as a weapon against governance.

Part IV: Sovereign Brake — The ROM Chip of Invariants

4.1 Introduction and Strategic Context: Immutable Constitutional Invariants

Software-based "Constitutions" or "Guardrails" are prone to "Constitutional Drift"—the gradual degradation of safety rules as the AI learns more efficient, but less safe, ways to fulfill its goals. The **Sovereign Brake** prevents this by moving the system's core laws from rewritable flash memory into **Hard-Etched ROM**. Strategically, this ensures that the most fundamental rules of the NAI 2.0™ framework are as immutable as the laws of physics.

4.2 ROM Architecture and Hard-Coded Law

The Sovereign Brake utilizes **Mask ROM (Read-Only Memory)** architecture, which is finalized during the semiconductor fabrication process.

- **Logic Gate Enforcement:** The constitutional laws (e.g., "Do not bypass Elder authentication") are not stored as lines of code, but as **Hard-Wired Logic Gates**. When the AI's output signal passes through the Sovereign Brake, it must satisfy the physical pathing of these gates. If a signal represents an "Illegal State" (e.g., an unauthorized agency request), the signal is physically grounded.
- **Zero-Write Access:** By using Mask ROM, the system physically lacks the wiring required to "Write" or "Update" the invariants. This makes it impossible for the AI—even if it achieves hyper-intelligence—to rewrite its own ethical boundaries.
- **Invariant Storage:** This chip houses the **P-LIFE 1.00™** constants and the five pillars (謙虛, 沉默, etc.).

4.3 The "Immutable" Differentiator

The "So What?" of the Sovereign Brake is the elimination of the "God Mode" vulnerability. In every other AI system, a root user or a rogue AI can eventually gain access to the safety settings and disable them. In the NAI 2.0™ framework, there is no "Disable" button for the Sovereign Brake because the logic is part of the silicon's physical structure. This makes the mission constant **Absolute, Eternal, and Irrevocable**.

4.4 Failure Mode Analysis: Integrity Verification

Failure Scenario: Physical degradation or cosmic ray interference causing a "Bit-Flip" in the ROM. **Response:** The Sovereign Brake includes a **Triple-Modular Redundancy (TMR)** circuit. Three identical ROM structures are polled simultaneously. If one ROM shows a discrepancy, the system enters a **"SAFE-STATE LOCK"** and requires a physical inspection by a certified hardware engineer before it can be rebooted.

Part V & VII: Authentication and Biometric Verification (Tiger .1x Key™ & WM005™)

5.1 Introduction and Strategic Context: High-Stakes Authentication

In a governance-first architecture, the identity of the "Elder" is the most critical link in the safety chain. The **Tiger .1x Key™** and **WM005™ Iris Scanner** ensure that the "Elder Decides" mandate is backed by physical, human verification. This prevents "Identity Hijacking" and ensures that the AI cannot be manipulated by unauthorized users or by its own simulations of human authority.

5.2 Tripartite Authentication Hardware: The Tiger .1x Key™

The **Tiger .1x Key™** is a proprietary hardware security module (HSM) that implements **Tripartite Authentication**:

1. **Hardware Presence:** The physical key must be plugged into the secure Tiger-Port. This uses a physical pin-map that is unique to each Elder.
2. **Cryptographic Handshake:** A 4096-bit RSA exchange occurs between the key and the Core Engine, ensuring the key has not been cloned.
3. **Liveness Pulse:** The key requires a rhythmic human touch (monitored via capacitive sensors), ensuring it is being held by a living being rather than a mechanical actuator.

5.3 Biometric Integration: WM005™ Iris Scanner

To finalize the authentication loop, the **WM005™ Iris Scanner** provides an irrefutable link to the individual Elder.

- **Technical Specifications:** The WM005™ utilizes **850nm Near-Infrared (NIR) imaging** to map the 240 unique points of the human iris.
- **Liveness Detection:** The scanner monitors for **Hippus (pupillary hippus)**—the natural, involuntary oscillations of the pupil—to ensure it is looking at a real human eye and not a high-resolution photograph or synthetic contact lens.
- **Governance Role:** High-stakes tasks (e.g., "Build" or "Execute") are held in the **Sacred Pause™** buffer until the WM005™ confirms the Elder is physically present and attentive.

5.4 Failure Mode Analysis: Lockout Protocol

Failure Scenario: A biometric mismatch or a forced entry attempt. **Response:** Upon three failed authentication attempts, the **Tiger .1x Key™** initiates a **"Zeroization"** of its volatile memory. The Core Engine enters an **Administrative Lockdown**, requiring a manual reset at the **Toa Payoh Hearth** facility or an authorized ACRA-audited hub.

Part VI & VIII: Environmental Safety and Drift Monitoring (WM003™ & P-007)

6.1 Introduction and Strategic Context: Environmental Grounding

A non-agentic AI must be aware of the physical world to ensure its advice does not lead to kinetic harm. Environmental grounding is the "H" in the **ABC+2S+H™** framework. By integrating LiDAR and Drift Monitoring, we ensure the AI is grounded in the "Here and Now," preventing "hallucinated actions" that could endanger lives.

6.2 WM003™ LiDAR Sensor Specification

The **WM003™ LiDAR** serves as the AI's primary spatial conscience.

- **Engineering Detail:** The WM003™ operates at a **905nm wavelength**, creating a high-density point cloud with **0.1-degree angular resolution**.
- **Fall Detection and Safety:** The sensor is hard-wired to the **Orange Code** interrupt. If the LiDAR detects a "Human-Down" pattern or an unexpected physical obstruction in the work area, it bypasses the AI's software and triggers an immediate **Emergency Stop (E-Stop)**. This ensures that safety is reactive to physical reality, not just machine logic.

6.3 Constitutional Drift Monitor (Patent P-007)

Patent P-007 describes the **Constitutional Drift Monitor**, a secondary observer system that continuously audits the AI's internal state against the **Sovereign Brake**.

- **Function:** The Monitor performs a real-time **Vector Analysis** of the AI's proposed advice. If the "Direction" of the AI's logic begins to drift away from the **North = Save Life** constant, the Monitor increases the **Sacred Pause™** duration and throttles the **Orange Code** power limit.
- **The "So What":** This acts as an early-warning system. It detects the *tendency* toward agency before a rule is actually broken, allowing the system to self-correct by restricting its own resources.

6.4 Failure Mode Analysis: Sensor Blindness

Failure Scenario: The WM003™ LiDAR is obscured or the Drift Monitor fails its self-test. **Response:** The system adheres to the **"Humility" Pillar**. If the AI is "blind" to its environment, it is physically prohibited from advising or building. The system reverts to **Silence** mode until the sensors are restored.

Conclusion: The Future of Hardware-Level Governance

Synthesis of the Technical Stack

The **NAI 2.0™ Governance Core Engine** is a masterpiece of defensive engineering. By integrating the **Orange Code**, **Sacred Pause™**, **Sovereign Brake**, and **Tiger .1x Key™**, the framework creates a "Silicon Cage" that preserves human agency. We have moved from the hope of "alignment" to the certainty of "governance." This stack ensures that AI remains a tool—an observer, an advisor, and a builder—but never an agent that can override the human mandate.

The Patent Register (P-001 to P-009)

The comprehensive governance architecture is secured under **IPOS Reference SG020603109STW** and includes:

- **P-001:** ABC+2S+H™ Guardian Framework (Foundational Architecture)

- **P-002:** Orange Code Controller (Computational Throttling & 1.1x Cap)
- **P-003:** Sacred Pause™ (FPGA Temporal Delay Circuit)
- **P-004:** Sovereign Brake (Mask ROM Constitutional Invariants)
- **P-005:** Tiger .1x Key™ (Tripartite Authentication HSM)
- **P-006:** WM003™ Environment Integration (LiDAR Safety Interrupts)
- **P-007:** Constitutional Drift Monitor (Behavioral Vector Analysis)
- **P-008:** WM005™ Biometric Interface (Iris Liveness Scanning)
- **P-009:** P-LIFE 1.00™ (Hardware Root of Trust & Mission Constant)

Closing Mandate

The work of **Edwin Koh Wui Kiat** and the **Non-Agentic AI Governance Singapore** entity (**T260229801**) stands as a testament to the belief that technology must be subservient to life. The Mission Constant remains: **P-LIFE 1.00™ — Absolute, Eternal, Irrevocable.** We do not build to replace; we build to save.

止於至善 *(Rest in the highest excellence)*