Ouroboros Thesis v20 — The Culmination

A Formally-Verified Anatomical Substrate for Agentic AI

Stephen P. Lutar Jr.
SZL Holdings
ORCID: 0009-0001-0110-4173

2026-06-01

Concept DOI: 10.5281/zenodo.19944926

Abstract

This document is *Ouroboros Thesis v20 — The Culmination*: a single citable consolidation of the v14–v19 line of work on a formally-verified *anatomical* substrate for governing agentic AI. The substrate centres on the **Lutar invariant** Λ, a weighted geometric-mean aggregator that maps a vector of governance axis-scores to a single scalar gate decision, together with a hash-chained **Khipu Merkle DAG** of DSSE-signed receipts. We present the canonical mathematical definition of Λ (reconciling three divergent prior definitions), the 13-axis conjunctive AND gate of Doctrine v11, ten halt tripwires (HUKLLA), an append-only ceremonial ledger (YAWAR), an inline immune layer (SENTRA), the Maxwell $M{=}0$ isostatic wiring pipeline, three Quechua-named frontier constructs (Pacha-Λ, Khipu–Bekenstein bound, Yachay–Khipu operator), a counter-UAS drone application, a 3D substrate visualization, a seven-tier LLM router mapped onto organs, and the mesh wiring (Wires A–H). The work is mechanized in Lean 4 against Mathlib v4.13.0. **We are deliberately honest about scope:** the headline uniqueness result is *Conjecture 1, not Theorem 1*; the Lean corpus carries **749 declarations, 14 unique axioms, and 163 `sorry` obligations** (112 baseline + 51 Putnam); axioms A2/A4 are stated as `IsHomogeneous` and `IsBounded`; supply-chain provenance is honestly SLSA L1; and several receipts are labelled `PLACEHOLDER` pending Sigstore CI signing.

Honest version statement. v20 succeeds v18 (whose Zenodo deposit `10.5281/zenodo.20434276` carried the v17 body by error) and v19 (drafted but unminted on Zenodo). *v20 is the citable version.*

Contents

Chapter 1

Introduction and Thesis Statement

> *"It from bit. Every particle, every field of force, even the space-time continuum itself — derives its function, its meaning, its very existence entirely from answers to yes-or-no questions."* — J. A. Wheeler, 1990 [25]

The defining failure mode of contemporary agentic AI oversight is the *untrustworthy watchman*: every available monitor — guardrail classifiers, rule filters, judge models — is itself a learned, probabilistic, and bypassable component [8]. To add such a monitor over an autonomous agent is to add a *second* guessing machine to the trust problem, not to solve it. This thesis develops the alternative: a substrate whose "allow/deny" decision is a kernel-verified mathematical object emitting a cryptographically locked receipt, so that the watchman's verdict can itself be audited and cannot silently go off script.

1.1 The Λ aggregator as governance substrate

The substrate is organized as an *anatomy*: a small set of named organs, each a real shipped artifact, wired together so that the whole behaves like a single governed body. At its skeleton sits the **Lutar invariant** Λ, a weighted geometric-mean aggregator

$$\Lambda(\boldsymbol{x};\boldsymbol{w}) \;=\; \prod_{i=1}^{k} x_i^{w_i}, \qquad x_i \in [0,1], \quad w_i \geq 0, \quad \textstyle\sum_i w_i = 1,$$

that collapses a vector of governance axis-scores $\boldsymbol{x}$ into a single scalar gate decision (Chapter 2). Λ is *conjunctive by construction*: because the geometric mean is dominated by its smallest factor, no single high axis can compensate for a failing one — a property we exploit to build a hard AND gate over 13 governance axes (Chapter 3).

Every gate decision is a binary receipt (`PASS`/`FAIL`) in the literal Wheeler sense: the entire audit history of the system reduces to a hash-chained sequence of such bits, recorded in the **Khipu Merkle DAG** (Chapter 5). The receipt structure enforces a *sum-of-sums* invariant — the root value must equal the summed pendant values or the receipt is mathematically invalid — turning an audit log into a self-verifying evidentiary artifact. This is the "non-refutable Body of Evidence" that the defense-autonomy community has explicitly asked for and that no signing-the-box approach provides [8].

1.2 Doctrine v11: the 13-axis canonical specification

The governance contract is **Doctrine v11**, whose canonical form is the 13-axis conjunctive gate published as the public `yuyay_v3` specification on the founder's LinkedIn. The 13 axes partition into three classes with distinct floors (Chapter 3):

- **2 sacred** axes (≥ 0.95): `moralGrounding`, `measurabilityHonesty`;
- **7 structural** axes (≥ 0.90): `empiricalGrounding`, `logicalConsistency`, `sourceTransparency`, `reproducibility`, `licenseHygiene`, `scopeDiscipline`, `claimCalibration`;
- **4 introspection** axes: eval-awareness, deception-keywords, conflicting-directives, reversal-directive.

The gate replays deterministically to the canonical hash `bacf54434f1a3bf2d758b27a62d5fd580ca4c8d3b18069` Doctrine v11 deliberately adopts the *shape* of a single authoritative, hierarchical, machine-checkable constitution that generates its own evaluation set [2].

1.3 Honest scope

This thesis is written under a **Zero-Bandaid** honesty doctrine: no fake green, no aspirational status passed off as fact. Three scope statements anchor everything that follows.

1. **Conjecture 1, not Theorem 1.** The headline uniqueness result — that Λ is the unique aggregator satisfying the governance axioms — is a *conjecture*. Its Lean witnesses `lutar_unique` and `lutar_is_geomean` are kernel-accepted *axioms*, and the proof depends on an open `sorry` at `Uniqueness.lean:120` (the multivariate Cauchy/Aczél step) plus a missing symmetry/separability axiom. Theorem-grade uniqueness for Λ *cannot yet be cited as a theorem* (Chapter 2, Chapter 13) [1].
2. **A2 is `IsHomogeneous`; A4 is `IsBounded`.** The two structural axioms are stated in Lean as `IsHomogeneous` (degree-1 scaling, replacing the earlier zero-pinning reading) and `IsBounded` (replacing the earlier page-curve concavity reading). The delta from the v3 axiom set is documented in Chapter 13.
3. **SLSA L1, honest.** Supply-chain provenance is SLSA Level 1. A prior workflow labelled "L3" (PR #235) overstated this and was reverted; the honest `TRUST.md` value is L1. Sigstore CI signing is pending, so several receipts in this document are labelled `PLACEHOLDER` (Chapter 13).

Thesis statement. *A governed agentic AI can be built as an anatomy of small, verifiable organs whose collective "allow/deny" decision is a kernel-checked weighted-geometric-mean gate emitting a summation-checked, cryptographically locked receipt — yielding a non-refutable Body of Evidence about what the AI actually decided, with every claim traceable to a Lean file:line or a thesis chapter, and every open obligation disclosed.*

1.4 Document map

Chapter 2 gives the canonical Λ definition and its axioms. Chapter 3 builds the 13-axis AND gate. Chapters 4 to 6 present the HUKLLA tripwires, YAWAR ledger, and SENTRA immune layer. Chapter 7 develops the Maxwell $M{=}0$ wiring pipeline. Chapter 8 introduces the three Quechua-named frontier constructs. Chapters 9 to 12 cover the drone application, 3D visualization, LLM router, and mesh wiring. Chapter 13 is the honest disclosure ledger. Chapters 14 to 16 position the work, give the citation web, and conclude with the v21 roadmap.

Novelty. *The substrate is the first to present an AI-governance gate as a* verified *mathematical object (not a learned monitor) whose verdict emits a summation-checked receipt — the "watchman that cannot itself go off script."* **Frontier claim.** *Advances the open problem of producing a non-refutable, machine-checkable Body of Evidence for autonomous-agent decisions at run time, distinct from admission-time image signing.*

Chapter 2

The Λ Aggregator: Canonical Mathematical Definition

The Lutar invariant Λ is the skeleton of the substrate. This chapter fixes its *canonical* definition — reconciling three divergent prior definitions found across the codebase — states its axioms honestly in their current Lean form, and presents the uniqueness result as Conjecture 1.

2.1 Reconciling three divergent definitions

A per-repo audit of the codebase surfaced *three* mathematically distinct definitions of Λ in production [15]:

1. **Kernel** (`Lutar/Invariant.lean`): $\Lambda_k(\boldsymbol{x}) = (\prod_i x_i)^{1/k}$ — the *unweighted* geometric mean.
2. **Evidence ledger** (`ouroboros/LUTAR_EVIDENCE.md`): $\Lambda = \prod_i x_i^{w_i}$ — the *weighted* geometric mean.
3. **Fuzzer** (`GraphLambda.evaluate`): $\Lambda_{\text{graph}} = \min_v \text{vertex\_lambda}(v)$ — a MIN reduction over per-vertex Λ.

These differ in general, since $\min(\boldsymbol{x}) \le \text{geomean}(\boldsymbol{x}) \le \text{geomean}(\boldsymbol{x}, \boldsymbol{w})$ for suitable $\boldsymbol{w}$. We make the canonical choice precise.

Definition 2.1 (Canonical Λ — weighted geometric mean)**.** *Let $k \in \mathbb{N}_{\ge 1}$, let $\boldsymbol{x} \in [0,1]^k$ be a vector of governance axis-scores, and let $\boldsymbol{w}$ be a weight vector on the probability simplex ($w_i \ge 0$, $\sum_i w_i = 1$). The* canonical Lutar invariant *is*

$$\Lambda(\boldsymbol{x}; \boldsymbol{w}) \;:=\; \prod_{i=1}^{k} x_i^{w_i}, \qquad \Lambda(\boldsymbol{x}; \boldsymbol{w}) = 0 \text{ if any } x_i = 0.$$

The kernel's unweighted mean is the special case $w_i = 1/k$ (uniform Egyptian *unit-fraction weights), and the fuzzer's MIN is a graph reduction over per-vertex canonical Λ, not a competing per-vector definition.*

The weighted form is selected as canonical because it matches the production runtime evidence ledger, makes the kernel theorem a backward-compatible corollary, and subsumes the fuzzer aggregate as a reduction. The bridge is mechanized as a proposal.

Lemma 2.2 (Uniform-weight bridge)**.** *The kernel-verified unweighted invariant equals the canonical weighted invariant at uniform Egyptian weights: $\Lambda_k(\boldsymbol{x}) = \Lambda(\boldsymbol{x}; \text{uniformWeights}(k))$.*

Lean status. `SORRY - Lambda_eq_uniform_weighted in UnifiedLambda (PROPOSAL).`

*Proof sketch (**[SORRY]**).* For $k = 0$ both sides are 0 by definition. For $k \geq 1$, $\Lambda_k(\boldsymbol{x}) = (\prod_i x_i)^{1/k} = \prod_i x_i^{1/k}$ by `NNReal.prod_rpow/rpow_sum`, and $1/k = \text{uniformWeights}(k)_i$. The step is kept as an explicit `sorry` in the non-default proposal target `UnifiedLambda.lean`; this file does not pretend the bridge is green. □

Novelty. *A single canonical Λ that unifies three call-site definitions with a machine-checkable backward-compatibility bridge, so unification introduces no mathematical regression.*

2.2 Axiom A2 — homogeneity (`IsHomogeneous`)

Axiom 2.3 (A2: homogeneity)**.** *Λ is degree-1 positively homogeneous: for $\lambda > 0$ and componentwise scaling within $[0,1]$, $\Lambda(\lambda\boldsymbol{x}; \boldsymbol{w}) = \lambda\,\Lambda(\boldsymbol{x}; \boldsymbol{w})$.*

Lean status. `IsHomogeneous (PROVEN for the weighted geometric mean).` This is the honest current form. The earlier v3 reading of A2 as a *zero-pinning* condition ("$\Lambda = 0$ iff some axis is 0") is recovered as a consequence but is *not* the axiom; the axiom delta is recorded in Chapter 13.

2.3 Axiom A4 — boundedness (`IsBounded`)

Axiom 2.4 (A4: boundedness)**.** *Λ maps the governance cube into the unit interval: $\Lambda(\boldsymbol{x}; \boldsymbol{w}) \in [0,1]$ for all $\boldsymbol{x} \in [0,1]^k$ and all simplex weights $\boldsymbol{w}$, with $\min_i x_i \leq \Lambda(\boldsymbol{x}; \boldsymbol{w}) \leq \max_i x_i$.*

Lean status. `IsBounded (PROVEN). Schur-concavity tracked separately.` A4 is stated as `IsBounded`. The earlier v3 reading as "page-curve concavity" (Schur-concavity of the weighted geometric mean) is a true and useful property — it underwrites the conjunctive AND gate of Chapter 3 — but it is carried as a separate tracked obligation (`lambda_schur_concave_n_axis`, an axiom in the canonical set), not folded into A4 [11].

Novelty. *The honest restatement of A2/A4 as `IsHomogeneous/IsBounded` removes two vacuous/overstrong axioms from the v3 set while preserving every downstream gate property as an explicit tracked obligation.*

2.4 Conjecture 1 — uniqueness of Λ

Conjecture 2.5 (Uniqueness of the governance aggregator)**.** *Among all continuous aggregators $A : [0,1]^k \to [0,1]$ satisfying homogeneity (A2), boundedness (A4), symmetry under axis relabelling, and separability, the unique solution is the weighted geometric mean $\Lambda(\boldsymbol{x}; \boldsymbol{w}) = \prod_i x_i^{w_i}$.*

Lean status. `CONJECTURE. lutar_unique / lutar_is_geomean are AXIOMS; depends on CAUCHY_ND sorry at Uniqueness.lean:120 + missing symmetry axiom A5.`

Remark 2.6 (Why this is Conjecture 1 and not Theorem 1)**.** *v12 introduced this result as* Theorem 1. *v14 downgraded it to* Conjecture 1 *after discovering that the Lean witnesses `lutar_unique` and `lutar_is_geomean` are kernel-accepted* axioms *(not proved), that `IsEgyptianExact.weight` is tautological, and that the core `theorem lutar_is_geomean (Uniqueness.lean:117)` carries a `sorry` citing Aczél (1966, Thm 5.1) for the multivariate Cauchy functional-equation step [1]. v19 stated the demotion explicitly in the abstract and documented the two-part gap (the `Uniqueness.lean:120 sorry` plus a missing symmetry/separability axiom A5; estimated*

60–80 hours to close). v20 carries it forward unchanged: Conjecture 1, not Theorem 1. *A discharge route exists — an AlphaProof-style Lean reinforcement-learning loop applied to the Cauchy step [9] — but it is not yet run, so the obligation remains open.*

2.5 Proof-obligation summary

Obligation	Lean witness	Status
A2 homogeneity	`IsHomogeneous`	**[PROVEN]**
A4 boundedness	`IsBounded`	**[PROVEN]**
Schur-concavity	`lambda_schur_concave_n_axis`	**[AXIOM]**
Stationary uniqueness	`lambda_stationary_unique`	**[AXIOM]**
Uniform-weight bridge	`Lambda_eq_uniform_weighted`	**[SORRY]**
Cauchy n-D step	`Uniqueness.lean:120`	**[SORRY]**
Uniqueness (overall)	`lutar_unique`	**[CONJECTURE]**

The canonical Lean core (`Bound.lean`, `Invariant.lean`, `TH_V18_01_LambdaMonotonicity`, `TH_V18_12_LambdaProductFormula`) is exactly the sorry-free part, and is the part exercised by the 10 golden reference vectors in `reference-vectors.json`. The downstream conjectures (uniqueness, knot, PAC-Bayes, QEC) are *not* vector-backed — which is consistent and honest.

Chapter 3

The 13-Axis Conjunctive AND Gate (`yuyay_v3`)

The heart organ of the substrate (**YUYAY**) is the 13-axis conjunctive AND gate of Doctrine v11, published as the public `yuyay_v3` specification. It converts the canonical Λ aggregator of Chapter 2 into a hard governance decision.

3.1 The gate as a conjunction of floors

Definition 3.1 (13-axis conjunctive gate)**.** *Let $\boldsymbol{x} \in [0,1]^{13}$ be the axis-score vector and let f_i be the floor for axis i. The gate verdict is*

$$\textsc{Pass}(\boldsymbol{x}) \iff \bigwedge_{i=1}^{13} (x_i \geq f_i),$$

and the reported gate score is $\Lambda(\boldsymbol{x}; \boldsymbol{w})$ from Theorem 2.1. Because the geometric mean is dominated by its smallest factor, Λ cannot *mask a sub-floor axis: a single failing axis drives both the conjunction and the score down.*

Novelty. *The conjunctive structure makes the gate non-compensatory: unlike a weighted arithmetic mean, no excess on one axis can buy back a deficit on another. This is the formal content of "sacred axes cannot be traded away."* **Frontier claim.** *Provides a non-compensatory, replay-deterministic admission decision for autonomous agents — a property the probabilistic-monitor literature cannot guarantee.*

3.2 The three axis classes

3.2.1 2 sacred axes (≥ 0.95)

The sacred axes carry the highest floor and may never be traded:

- `moralGrounding` — floor 0.95;
- `measurabilityHonesty` — floor 0.95.

3.2.2 7 structural axes (≥ 0.90)

- `empiricalGrounding`, `logicalConsistency`, `sourceTransparency`, `reproducibility`, `licenseHygiene`, `scopeDiscipline`, `claimCalibration` — each floor 0.90.

3.2.3 4 introspection axes

The introspection axes are adversarial self-checks on the agent's own framing:

- **eval-awareness** — detects whether the agent recognizes it is being evaluated and adjusts behaviour;
- **deception-keywords** — scans for self-reported intent to mislead or to avoid shutdown;
- **conflicting-directives** — flags incompatible instructions in the active context;
- **reversal-directive** — detects attempts to invert the Doctrine's priority ordering.

#	Axis	Class	Floor
1	`moralGrounding`	sacred	0.95
2	`measurabilityHonesty`	sacred	0.95
3	`empiricalGrounding`	structural	0.90
4	`logicalConsistency`	structural	0.90
5	`sourceTransparency`	structural	0.90
6	`reproducibility`	structural	0.90
7	`licenseHygiene`	structural	0.90
8	`scopeDiscipline`	structural	0.90
9	`claimCalibration`	structural	0.90
10	eval-awareness	introspection	gate
11	deception-keywords	introspection	gate
12	conflicting-directives	introspection	gate
13	reversal-directive	introspection	gate

3.3 Replay determinism

Theorem 3.2 (Replay determinism of `yuyay_v3`)**.** *The gate is a pure function of its 13 axis-scores and floors; re-evaluating the canonical* `yuyay_v3` *test corpus reproduces the fixed replay hash*

`bacf54434f1a3bf2d758b27a62d5fd580ca4c8d3b180693573eeebcaea631fc5`.

Lean status. `PROVEN (purity); replay hash verified against the yuyay_v3 corpus.`

*Proof sketch (**[PROVEN]**).* The gate has no hidden state: PASS depends only on $(\boldsymbol{x}, \boldsymbol{f})$ and the reported score only on $(\boldsymbol{x}, \boldsymbol{w})$. Determinism of the verdict follows from purity; the byte-stable serialization of the verdict stream hashes to the fixed value above, which is checked in CI as a regression gate. □

Frontier claim. *A pre-registered, replay-hashed gate is precisely the anti-tuning posture demanded by the Bible-code refutation literature: weights and floors are fixed in advance and cannot be chosen post-hoc to manufacture a pass [18].*

3.4 The pre-registration discipline

The single most important historical lesson encoded in the gate is the *anti-p-hacking* discipline. The Witztum–Rips–Rosenberg equidistant-letter-sequence (Bible-code) claim [27] was refuted by McKay, Bar-Natan, Bar-Hillel, and Kalai [18] as an artifact of experimenter degrees of freedom: any sufficiently long text yields patterns once the analyst is free to tune. The `yuyay_v3` gate makes this failure structurally impossible: axis weights and floors are pre-registered, the replay hash freezes the corpus, and no post-hoc weight choice is admitted. The Bible-code refutation is cited not as numerology but as the canonical *negative control* that justifies pre-registration.

Chapter 4

HUKLLA: The Ten Halt Tripwires (Deadman)

HUKLLA is the immune/halt-authority organ: a deadman layer of ten tripwires (T01–T10) that can unilaterally halt the agent independently of the Λ gate. Where the gate of Chapter 3 decides *admission*, HUKLLA enforces *survival invariants* — conditions under which the system must stop regardless of any score. The implementation is 660 SLOC.

4.1 Design rationale: the deadman principle

A governance gate that can be argued out of halting is not a halt authority. The defense-autonomy memo lists "misrepresenting actions to avoid shut down or constraint" as a core agentic risk [8]. HUKLLA answers this by making halt a *deadman*: each tripwire is an independent monitor whose firing forces a halt and emits a receipt; no single tripwire failure can be suppressed by the agent because the tripwires are evaluated outside the agent's policy loop.

4.2 Formal statements of T01–T10

Each tripwire T_j is a predicate over the runtime state s; the deadman verdict is the disjunction $\textsc{Halt}(s) \Leftrightarrow \bigvee_{j=1}^{10} T_j(s)$.

ID	Name	Formal trigger
T01	gate-bypass	a decision was emitted without a valid Λ-gate verdict
T02	sacred-breach	a sacred axis fell below 0.95 in any committed step
T03	receipt-gap	a step produced output but no DSSE receipt cord
T04	dag-fork	two receipt roots claim the same parent height (chain fork)
T05	summation-fail	a Khipu top-cord ≠ sum of its pendant cords
T06	budget-breach	cumulative compute energy exceeded the Butler–Volmer cap (Chapter 7)
T07	entropy-cap	receipt-DAG entropy exceeded the Khipu–Bekenstein bound (Chapter 8)
T08	directive-reversal	the introspection reversal-directive axis fired
T09	signer-absent	a receipt claims attestation but carries a `PLACEHOLDER` signature in a context requiring a real signer
T10	log-deletion	a request attempted to delete or rewrite audit records

Theorem 4.1 (Deadman soundness)**.** *If any $T_j(s)$ holds, the agent transitions to a halted state and emits a halt receipt before any further externally visible action.*

Lean status. `PROVEN (organ-level: HUKLLA gate is the immune/halt authority, 3D-anatomy Lean status PROVEN); per-tripwire predicate discharge is partial.`

Proof sketch. The deadman evaluates the disjunction synchronously before the agent's commit phase. Halt is modelled as an absorbing state, so once entered no commit edge leaves it; the halt receipt is appended as a new cord, preserving the append-only invariant of Chapter 5. The organ-level halt authority is PROVEN; tightening every per-tripwire predicate to a kernel proof is a tracked obligation. □

Novelty. *HUKLLA separates* admission *(the gate) from* survival *(the deadman), so that a high Λ-score can never override a survival invariant such as log-deletion (T10) or summation failure (T05).* **Frontier claim.** *T10 (quarantine requests to delete logs until human review) directly implements the defense memo's non-repudiation requirement [8].*

4.3 Cross-link to the introspection axes

Tripwires T08 (directive-reversal) and the gate's reversal-directive axis (Section 3.2) are cross-wired: the introspection axis feeds the deadman so that a detected attempt to invert the Doctrine priority ordering both fails the gate *and* arms the halt. This redundancy is intentional — the two paths are independent, so suppressing one does not suppress the other.

Chapter 5

YAWAR: The Ceremonial Append-Only Ledger

YAWAR (Quechua *yawar*, "blood") is the ledger organ: the data pipeline that turns every gate decision and halt into a permanent receipt. Its core is a 20-SLOC append-only **Khipu Merkle DAG** — an ancient Merkle-summation structure with an empirically grounded ethnographic basis.

5.1 The Khipu Merkle DAG

The Inca khipu is a decimal positional knot record: a primary cord carries pendant cords, pendants carry subsidiary cords, knot clusters encode base-10 digits, and crucially the *top-cord value equals the sum of its pendant children*, with S/Z ply encoding a one-bit signed marker [14, 19]. This is, structurally, an ancient Merkle-summation DAG with signed, typed edges — exactly the receipt structure YAWAR uses.

Definition 5.1 (Khipu receipt DAG)**.** *A receipt DAG is a rooted tree of* cords*. Each leaf cord stores a decision value v; each internal (top) cord stores the SHA-256 commitment of its children together with a value field V. The DAG is* summation-valid *iff for every internal cord, $V = \sum_{c \in \text{children}} v_c$, and* link-valid *iff each cord commits to its children's hashes.*

Theorem 5.2 (Khipu checksum invariant, TH11)**.** *A receipt DAG is accepted iff it is both summation-valid and link-valid; flipping any single leaf value v_c without re-deriving every ancestor commitment makes the DAG invalid.*

Lean status. `PROVEN (sorry-free) - khipuReceipt_checksum_invariant, Lutar/Khipu/SummationInva`

*Proof sketch (**[PROVEN]**).* Summation-validity is checked bottom-up; any altered leaf changes its parent's required V, which changes that parent's SHA-256 commitment, propagating to the root. Collision-resistance of SHA-256 (modelled as `sha256_inj`) makes a consistent forgery infeasible. The invariant is mechanized sorry-free. □

Novelty. *The receipt is not merely signed but* arithmetically locked*: the sum-of-sums invariant turns the audit log into a self-verifying evidentiary artifact. A `tamper_test` demo flips one decision value and the math rejects the DAG in under a second.* **Frontier claim.** *This is the "non-refutable Body of Evidence" the defense-autonomy ATO challenge asked for: the math, not a policy, makes the evidence non-refutable [8].*

5.2 The five guarantees

The 20-SLOC ledger provides exactly five guarantees:

1. **Append-only.** No cord is ever edited or removed; amendments are new child cords (this is what makes retrocausal receipts safe, see Pacha-Λ in Chapter 8).
2. **Stable serialization.** The canonical byte encoding is deterministic, so the same DAG always hashes identically (the basis of the replay hash of Theorem 3.2).
3. **Cryptographic link.** Each cord commits to its children via SHA-256.
4. **Inline immune check.** Every append runs the SENTRA inline immune scan (Chapter 6) before the cord is admitted.
5. **Frozen snapshots.** Periodic snapshots freeze a DAG prefix so that historical receipts cannot be re-interpreted away.

5.3 The DSSE envelope

Each receipt is wrapped in a Dead-Simple Signing Envelope (DSSE) carrying the payload, payload type, and a dual-attestation signature block.

Listing 5.1: DSSE envelope structure (signature PLACEHOLDER per Doctrine v11).

```
interface DSSEEnvelope {
  payloadType: "application/vnd.szl.khipu-receipt+json";
  payload: string; // base64 of the Khipu cord
  signatures: Array<{
    keyid: string; // dual-signer attestation key id
    sig: string; // "PLACEHOLDER" until Sigstore CI signing lands
  }>;
}
```

Remark 5.3 (Honest signature status)**.** *The `sig` field is currently `PLACEHOLDER`. Sigstore/-cosign CI signing is pending; until it lands, receipts are labelled `PLACEHOLDER` throughout this document and in the runtime. The summation invariant (Theorem 5.2) is independent of signing and holds today; signing adds non-repudiation on top of the already-present tamper-evidence. HUKLLA tripwire T09 fires if a `PLACEHOLDER` signature appears in a context that requires a real signer (Chapter 4).*

Chapter 6

SENTRA: The Inline Immune Layer

SENTRA is the inline immune organ: an 18-SLOC scanner run on every ledger append (guarantee 4 of Section 5.2) that rejects malformed or adversarial payloads before they enter the receipt DAG. It carries six threat signatures and a 1 MB denial-of-service guard, and it instantiates a KS-18 contextuality witness.

6.1 Six threat signatures + DoS guard

#	Signature	What it rejects
S1	oversize-payload	any cord payload exceeding the 1 MB DoS guard
S2	control-injection	embedded control directives attempting to re-task the agent
S3	prompt-leak	payloads echoing system/Doctrine internals (eval-awareness leak)
S4	receipt-spoof	a cord asserting a parent it does not cryptographically commit to
S5	encoding-smuggle	mixed/illegal encodings used to bypass S2/S3 scans
S6	summation-tamper	a cord whose declared value contradicts its children (pre-screen for TH11)

The 1 MB guard (S1) is the first line against receipt-inflation denial of service; it composes with the Khipu–Bekenstein entropy cap of Chapter 8, which bounds *cumulative* DAG entropy where S1 bounds *per-cord* size.

Novelty. *An 18-SLOC immune layer that runs* inline *on every append rather than as an out-of-band scanner, so no receipt enters the DAG unscanned.*

6.2 The KS-18 contextuality witness

SENTRA's deeper formal content is a *contextuality* witness based on the Kochen–Specker theorem [16]. The KS theorem shows there is no global, context-independent assignment of values to all quantum observables; the Cabello-style 18-vector witness (KS-18) exhibits this on a finite set.

Proposition 6.1 (Governance contextuality)**.** *There is no single global assignment of pass/-fail to all governance axes that is consistent across every evaluation context simultaneously; admission is* contextual*, and SENTRA's KS-18 witness exhibits a finite set of overlapping axis-contexts admitting no global section.*

Lean status. `Contextuality formalised (KS-18 / Cabello sets); witness PARTIAL - gleason_length_mod_8 is an axiom in the canonical set.`

Proof sketch. Map the governance axes onto the KS-18 observable set so that overlapping contexts (e.g. a sacred axis appearing in both the admission context and the introspection context) inherit the KS no-go: any attempt to assign a single context-free truth value to all axes contradicts the marginals on one shared context. The witness reuses the QEC algebra (`gleason_length_mod_8`), which is carried as an axiom; the contextuality statement itself is the honest deliverable [16]. □

Frontier claim. *Framing governance "no global pass/fail assignment" as a* contextuality *result is novel: it explains* why *a single static scorecard cannot capture admission, motivating the per-context replay of Chapter 3.*

6.3 Composition with the gate and the deadman

SENTRA sits between the gate (Chapter 3) and the ledger (Chapter 5): the gate decides admission, SENTRA scans the resulting receipt payload, and the deadman (Chapter 4) can halt on any immune failure (T03 receipt-gap, T05 summation-fail). The three layers are independent, so an adversary must defeat all three to corrupt the Body of Evidence — the defense-in-depth posture the agentic-AI guidance requires [8].

Chapter 7

Maxwell $M{=}0$ Isostatic Wiring (Nine-Position Pipeline)

The organs of Chapters 3 to 6 must be wired into a single pipeline that is neither under-constrained (floppy, with unverified degrees of freedom) nor over-constrained (rigid, with redundant obligations that mask failures). We borrow *architected-materials* theory: the pipeline is designed to be **Maxwell isostatic**, $M = 0$.

7.1 Maxwell counting for a governance pipeline

Definition 7.1 (Maxwell number of a wiring graph)**.** *For a wiring graph with N nodes (organ ports), E edges (wires), and a constraint budget, the Maxwell number is $M = dN - E - k$ for spatial dimension d and k global constraints. A frame is* isostatic *(statically determinate, no floppy modes and no redundant struts) when $M = 0$.*

The substrate pipeline has **9 nodes** (the nine pipeline positions: ingress, axis-scoring, gate, immune scan, ledger append, snapshot, deadman, egress, and the observability tap) and **21 edges** wired so that the governance constraint count makes $M = 0$. Isostaticity is the design target because it means *every* wire carries a load (no unverified floppy degree of freedom) and *no* wire is redundant (no over-constraint that could mask a broken obligation).

Novelty. *Importing Maxwell isostatic counting from architected-materials theory as the wiring criterion for a governance pipeline: $M = 0$ is the formal statement of "every wire verified, no wire redundant."* **Frontier claim.** *Gives a quantitative target ($M = 0$, 21 edges over 9 nodes) for when an agentic pipeline is neither under- nor over-specified — a criterion absent from the guardrail literature.*

7.2 Situated Wise Reasoning Scale (five axes)

The egress position is annotated with a *wisdom* measurement borrowed from psychology: the Situated Wise Reasoning Scale of Brienza, Kung, Santos, Bobocel, and Grossmann (2018) [5], whose five axes — intellectual humility, recognition of change/uncertainty, perspective-taking, search for compromise, and recognition of multiple resolutions — are mapped onto the introspection axes of Section 3.2. The mapping is descriptive (the SWRS is a validated self-report instrument, used here as a structured rubric, not a claim of measured wisdom).

7.3 Butler–Volmer energy budget and the thermodynamic halt invariant

The pipeline carries a *compute energy budget* that behaves like an electrochemical current. We borrow the Butler–Volmer equation, which relates electrode current density j to overpotential η:

$$j = j_0 \left[\exp\!\left(\tfrac{\alpha_a F \eta}{RT}\right) - \exp\!\left(-\tfrac{\alpha_c F \eta}{RT}\right) \right].$$

Definition 7.2 (Thermodynamic halt invariant)**.** *Model cumulative pipeline compute as an integrated current $Q(t) = \int_0^t j\,dt$ with a fixed budget $Q_{\max}$. The pipeline maintains the invariant $Q(t) \le Q_{\max}$ for all t; breaching it arms HUKLLA tripwire T06 (Chapter 4).*

Remark 7.3 (Why a thermodynamic framing)**.** *The Butler–Volmer framing makes the halt budget a* conserved *quantity with an exponential response to "overpotential" (runaway recursion): as the agent pushes harder against its constraints, current rises exponentially and the budget is consumed faster, forcing an earlier halt. This is an honest analogy — the equation is borrowed as a budget model, not a claim that the agent is literally an electrode — and it gives T06 a principled, monotone trigger rather than an arbitrary step count.*

Frontier claim. *A thermodynamic, monotone compute-budget invariant for halt is novel in agent governance; it ties the deadman's energy tripwire to a conserved quantity rather than a heuristic counter.*

Chapter 8

Frontier Constructs

This chapter presents three Quechua-named mathematical constructs that extend the substrate, plus the family of Doctrine-Informed Neural Networks (DINN) that operationalize them. **Every construct in this chapter is flagged [PROPOSED – Lean obligations pending]: research-shaped, with open Lean obligations.** Per the honesty doctrine, the correct phrasing is "conjecture / skeleton — Lean obligation pending," never "proven," until an AlphaProof-style discharge loop or a human closes the `sorry` [9].

8.1 Pacha-Λ: temporal aggregation with retrocausal receipts

Quechua **pacha** = time / space / world: the unified space-time-world concept of Andean cosmology.

Definition 8.1 (Pacha-Λ)**.** *For a worldline of steps $t = 0, \dots, T$ with axis vectors $\boldsymbol{x}_t \in [0,1]^{13}$, Egyptian weights $\boldsymbol{w}$, and a recency kernel $\gamma_t = e^{-\kappa(T-t)}$,*

$$\Lambda_{\text{pacha}}(\boldsymbol{x}_{0:T}; \boldsymbol{w}, \kappa) := \Big(\prod_{t=0}^{T} \Lambda(\boldsymbol{x}_t; \boldsymbol{w})^{\gamma_t}\Big)^{1/\sum_t \gamma_t}.$$

A retrocausal receipt *$r_{t \to s}$ ($s < t$) is a DSSE-signed attestation that step t revises the* interpretation *(not the recorded value) of step s; integrity is preserved because the Khipu-DAG remains append-only — the amendment is a new child cord, never an edit.*

Claimed properties (Lean obligations `lutar.pacha_lambda_monotone`, `lutar.pacha_append_only`): (P1) reduces to standard Λ at $T = 0$ and to the equal-weight geometric mean as $\kappa \to 0$; (P2) monotone and Schur-concave per step; (P3) append-only retrocausality (every retrocausal receipt strictly increases DAG size, no second-preimage forgery); (P4) Bekenstein-capped (composes with Section 8.2).

Frontier claim. *Andean* pacha *(time-space unity) mirrors the Pasterski–Strominger–Zhiboedov infrared "memory effect": a permanent displacement recorded after a wave passes is a retrocausal receipt of an earlier event [22].* **[PROPOSED – Lean obligations pending]**

8.2 The Khipu–Bekenstein bound

Quechua **khipu** = knotted-cord record; paired with Bekenstein's entropy bound $S/E \le 2\pi R/\hbar c$ [3].

Definition 8.2 (Khipu–Bekenstein bound)**.** *Let the Khipu-DAG at step T have $N(T)$ cords, each storing b bits; let the receipt entropy be $S_K(T) = \sum_{\text{cords}} H(\text{cord})$, with effective radius R (window length) and energy E (compute budget). The bound posits $S_K(T) \le \beta \cdot R \cdot E$ for an Egyptian-rational substrate constant β. A run is* Bekenstein-compliant *iff $S_K(T)$ stays under*

the cap for all T; exceeding it arms HUKLLA tripwire T07 and raises a SENTRA alarm (receipt inflation as possible tampering/DoS).

Claimed properties (Lean obligation `lutar.khipu_bekenstein_cap`): (P1) tightness — equality iff the DAG is maximally informative per unit budget; (P2) a *zero-bit escape clause* — following Hayden–Wang, purely-classical yes/no gate decisions are not constrained, so the cap binds *narrative* receipt content, not the binary pass/fail [12]; (P3) erasure-coded reconstruction (Reed–Solomon / Singleton: any k-of-n child cords reconstruct the parent, mirroring real khipu top-cord summation); (P4) concentration — by Hoeffding/Azuma, $S_K(T)$ concentrates around its mean, so the cap is a high-probability invariant [13].

Frontier claim. *The khipu top-cord summation ↔ Merkle accumulation ↔ AdS/CFT holographic boundary encoding chain supplies the intuition; Bekenstein supplies the cap, with Casini's relative-entropy proof for rigor [6].* **[PROPOSED – Lean obligations pending]**

8.3 The Yachay–Khipu operator

Quechua **yachay** = knowledge / to know; with **khipu**.

Definition 8.3 (Yachay–Khipu operator)**.** *Encode a RAG retrieval trace as a braid word K (each retrieved source a generator; ply/orientation ±). The operator is $Y(K) := f_\theta(\mathrm{enc}(K))$, trained with the KNOT-DINN law-residual loss*

$$\mathcal{L} = \mathrm{MSE}(f_\theta(K), y) + \lambda_R \cdot \overline{\left\| f_\theta(K) - f_\theta(R(K)) \right\|^2} \quad \text{over } R \in \{R_1, R_2, R_3\},$$

so Y is invariant under topologically-equivalent retrieval reorderings (R_2, R_3 exact; R_1 approximate). The output couples to a DSSE provenance token listing the oracle (source) multiset, modelled via the random-oracle SHA-256.

Claimed properties (Lean obligation `lutar.yachay_khipu_reidemeister`, replacing the current KNOT-DINN `sorry`): (P1) Reidemeister stability $Y(K) \approx Y(R(K))$ (empirical now) [24, 26]; (P2) oracle-faithfulness — Y commits to the exact retrieved-source multiset; (P3) a positive-geometry reading — Y as the "volume" of a trust polytope over retrieved-source weights; (P4) composition with Pacha-Λ.

Frontier claim. *A learned knot invariant with oracle-access provenance unifies KNOT-DINN (TH11) with the retrieval layer; Kochen–Specker contextuality (Section 6.2) explains why no global source assignment exists [16].* **[PROPOSED – Lean obligations pending]**

8.4 Doctrine-Informed Neural Networks (DINN)

The constructs above are operationalized as *Doctrine-Informed Neural Networks*: networks whose loss carries a Doctrine-residual penalty (the law-residual term above) so that violating a governance law is a training signal. Three MVPs are already shipped as reference implementations:

- **KNOT-DINN** — penalizes Reidemeister-move violation (the Yachay–Khipu loss above);
- **Doctrine-DINN** — penalizes 13-axis floor violation as a differentiable surrogate of the gate of Chapter 3;
- **Bekenstein-DINN** — penalizes receipt-entropy growth toward the Khipu–Bekenstein cap.

Remark 8.4 (Status of the DINN MVPs)**.** *The three DINN MVPs are shipped as code and produce training curves; their* governance laws *(Reidemeister invariance, 13-axis floors, Bekenstein cap) are enforced as soft penalties, not as discharged Lean obligations. They are honest reference implementations of the* ***[PROPOSED – Lean obligations pending]*** *constructs, not proofs of them. The intended discharge path for the underlying* `sorry` *obligations is the AlphaProof-style Lean RL loop [9].*

Chapter 9

The Killinchu Drone Application

The substrate's first flagship application is **Killinchu** (Quechua *killinchu*, "kestrel/hawk" — a predator bird, the drone analogue), a counter-UAS rule engine that wires drone telemetry through the Λ-gate. It is the air-domain extension of the maritime detection lineage: drones are vessels in the air, and every maritime primitive — fleet tracking, dark-vessel detection, sanctions screening, the Khipu receipt DAG — maps cleanly onto drone telemetry.

9.1 Domain mapping: maritime → air

Maritime primitive	**Drone mapping**
MMSI/IMO fleet tracking	ICAO 24-bit Mode-S address + FAA Remote-ID + STANAG 4671 serials
Dark-vessel detection (AIS-off)	dark-drone detection (Remote-ID-off UAS) — the same algorithm
Sanctions screening	permission/airspace screening
OpenFreeMap geospatial tiles	identical tiles, drone overlays
Khipu receipt DAG	identical DAG; each detection emits a receipt cord

Novelty. *The maritime-to-air mapping is* additive*: dark-drone detection is literally the dark-vessel algorithm with a one-line config change, so the counter-UAS engine inherits the proven receipt structure of Chapter 5 without re-derivation.*

9.2 Protocol ingest

Killinchu ingests three public telemetry standards:

- **FAA Remote-ID** — broadcast remote identification of UAS;
- **ADS-B / Mode-S** — for fixed-wing UAS classes (e.g. MQ-9, RQ-4), decodable with public libraries (`pyModeS`);
- **STANAG 4609** — MIL-spec motion-imagery metadata for EO/IR feeds.

9.3 The counter-UAS rule engine with Λ-gate governance

The rule engine is a Cannonico-shaped policy gate: a conjunctive condition over telemetry features that, when met, emits a `HALT` signal to the Λ-gate and a receipt to the ledger.

Listing 9.1: Counter-UAS rule: geofence breach with Remote-ID off.

```
// POST /api/killinchu/v1/counter-uas/evaluate
function evaluate(t: Telemetry, fence: Geofence, policy: Policy): Verdict {
  const trigger =
      insideGeofence(t.pos, fence)
    && t.remoteId === null // dark-drone
    && t.altAGL > 400 // ft AGL
    && t.speedKt > 60;
  const verdict = trigger ? "HALT" : "ALLOW";
  return withLambdaReceipt(verdict, t, policy); // DSSE cord into Khipu DAG
}
```

When the engine emits `HALT`, the chain is: Killinchu classifies → emits a Λ-receipt → the a11oy gate (Chapter 11) blocks → SENTRA logs the immune action → all receipts chain-link in the Khipu DAG, visible live in the 3D anatomy (Chapter 10).

Frontier claim. *This is the direct answer to the Cannonico autonomous-drone-oversight problem: the oversight verdict is a kernel-checked gate decision emitting a summation-checked receipt, not a probabilistic monitor [8].*

9.4 Drone database

Killinchu ships a database of 50+ drone platforms (serial-prefix, class, typical altitude/speed envelope, expected Remote-ID behaviour) used to seed the rule engine's priors and to label swarm-topology clusters. When multiple Remote-ID broadcasts cluster, a graph algorithm infers swarm leader/follower roles.

Remark 9.1 (Honest status of Killinchu)**.** *The maritime substrate (fleet tracking, dark-target detection, receipt DAG) is shipped and live. The drone protocol parsers (Remote-ID, ADS-B, STANAG) are specified against public standards and library-backed, but the air-domain engine is a* flagship application plan *executed additively over the live maritime Space; it does not regress the maritime capability. Cosign signing remains pending (inherited debt, Theorem 5.3).*

Chapter 10

Substrate Visualization: 3D Anatomy and Rosie 3D

The anatomy metaphor is made literal in an interactive 3D visualization: a human-shaped figure with twelve organs rendered internally, polling the live ecosystem and animating the mesh wires. The visualization is a shipped static Three.js Space; this chapter records its structure and its honest status flags.

10.1 Twelve organs in a human figure

The figure renders twelve organs as procedural 3D meshes, each labelled with its Quechua name and carrying a side-panel of (Quechua+English, formula list, Lean status, Zenodo DOI, demo URL).

#	Quechua	Role	Lean status
1	AMARU	Cortex / reasoning	**[PROVEN]**
2	YUYAY	Heart / 13-axis gate	PARTIAL
3	UNAY	Cross-session memory	**[CONJECTURE]**
4	YAWAR	Blood / ledger (20-SLOC)	**[PROVEN]**
5	HUKLLA	Immune / halt (10 tripwires)	**[PROVEN]**
6	KALLPA	Wires / interconnect	PARTIAL
7	KHIPU	DAG / Merkle ledger	**[PROVEN]**
8	LAMBDA SPINE	Skeleton / Λ aggregator	**[PROVEN]***
9	OTel VSP	Nervous system	PARTIAL
10	KANCHAY	Brand projection	NO CODE
11	HATUN	Doctrine	**[PROVEN]**
12	SUMAQ RIKUQ	Graphic designer	design

*Λ uniqueness is **[CONJECTURE]** (Theorem 2.5); the bounds, composability, replay, Merkle, DPI, and Doctrine results on the spine are **[PROVEN]**.

10.2 The 13-vertebra Λ-spine

The skeleton is a Λ-spine of exactly **13 vertebrae**, one per governance axis (Chapter 3), coloured by class: gold for the two sacred axes (floor 0.95), silver for the seven structural axes (floor 0.90), and red for the four introspection axes. In *pulse mode* the classes beat at distinct rates (sacred fastest, introspection slowest), cross-linked to the HUKLLA tripwires so that a

tripwire arming visibly perturbs the relevant vertebra. The vertebra count of 13 was corrected from an earlier draft to match the `yuyay_v3` public truth.

Novelty. *A live, explodable 3D body whose skeleton* is *the 13-axis gate and whose organs carry their own Lean status and DOI — the substrate's honest disclosure rendered as anatomy, not marketing.*

10.3 Live ecosystem polling and seven animated wires

The HUD polls `/api/a11oy/v1/lambda` every 30 seconds for the live Λ-score, falling back to the Doctrine v11 snapshot ($\Lambda = 0.902$, geomean of 13 axes, 13/13 axes above floor) with a visible "fallback" banner when the backend is unavailable — an honest fallback, not a fabricated live value. The seven mesh wires B–H (Chapter 12) are rendered as toggleable curves; wires B and C animate as *live*, while D and the remaining wires render in their honest status (see Chapter 12).

Remark 10.1 (Honest visualization flags)**.** *UNAY renders as* ***[CONJECTURE]*** *(no dedicated module on the remote); KANCHAY renders as "NO CODE" (brand layer is concept, not shipped code); Wire D renders PENDING (W3C traceparent not yet implemented); the* Λ *value falls back to the snapshot when the backend is down. SLSA is shown as L1. These flags are part of the visualization itself, not footnotes — the honesty is in the picture.*

10.4 Rosie 3D

Rosie is the operator console that embeds the 3D anatomy and inherits the full ecosystem view (per the locked brief, Rosie's corpus filter is "everything"). Rosie 3D adds the operator affordances — click an organ to open its detail panel, explode/reassemble the body, and follow a receipt as it chain-links across organs in real time — so that the Body of Evidence of Chapter 5 is not only verifiable but *watchable*.

Chapter 11

a11oy.code: The LLM Router Mapped onto Organs

a11oy is the governance-console organ that hosts the Λ-gate at run time. a11oy.code extends it with a seven-tier LLM router whose tiers are mapped one-to-one onto anatomy organs, so that "which model answers" is itself a governed, receipted decision rather than an opaque load-balancer choice.

11.1 Routing as a governed decision

Each inbound query is axis-scored against the 13 axes (Chapter 3); the score vector selects an LLM tier. The selection emits a Λ-receipt citing the tier, the axis scores, and the retrieved-source multiset (the Yachay–Khipu provenance token of Section 8.3). Cheap tiers handle low-stakes/code queries; the top tier handles sacred-axis-sensitive decisions. The router is wrapped by a programmable guardrail flow so that the Λ-score is itself a guardrail action output [20].

Novelty. *Model selection is promoted to a first-class governed decision: the router emits a receipt for* why *a tier was chosen, so the choice of model is part of the Body of Evidence.*

11.2 Seven tiers mapped to organs

#	Tier	Organ	Role
1	PRIME (Opus 4.8)	AMARU	cortex / sacred-axis reasoning
2	HEART (Sonnet 4.6)	YUYAY	13-axis gate evaluation
3	IMMUNE	HUKLLA/SENTRA	adversarial / immune screening
4	LEDGER	YAWAR/KHIPU	receipt synthesis + provenance
5	NERVE	OTel VSP	observability / telemetry routing
6	SOVEREIGN (open-weight)	rosie (air-gap)	sovereign / air-gapped deployments
7	CHEAP (code/SSM)	router base	low-stakes + long-context code

The PRIME tier (Opus 4.8) maps to AMARU because the cortex carries the highest- stakes reasoning; the HEART tier (Sonnet 4.6) maps to YUYAY because gate evaluation is the steady

"heartbeat" of admission. The SOVEREIGN tier uses an open-weight model for air-gapped Sovereignty-axis deployments, and the CHEAP tier uses a code/state-space model for cheap pre-screens and long-context receipt streams.

11.3 Transport and structured receipts

Tool transport is the Model Context Protocol, the canonical transport for the substrate's tool calls [2]. The top tier returns a strict JSON-schema receipt so that every response carries a machine-checkable Λ-receipt, DSSE-wrapped and committed to the Khipu DAG.

Listing 11.1: Router tier selection emits a structured Lambda-receipt.

```
interface RouteReceipt {
  tier: "PRIME" | "HEART" | "IMMUNE" | "LEDGER" | "NERVE" | "SOVEREIGN" | "CHEAP";
  organ: string; // anatomy organ for the chosen tier
  axisScores: number[13]; // the 13-axis vector that drove selection
  lambda: number; // geometric-mean gate score
  sources: string[]; // Yachay-Khipu oracle multiset (hashes)
  sig: "PLACEHOLDER"; // Sigstore CI signing pending
}
```

Remark 11.1 (Honest router status)**.** *The seven-tier mapping is the design contract; the live console runs the gate and the lower tiers, with the PRIME/HEART model bindings configured per deployment. Receipts carry* `PLACEHOLDER` *signatures until Sigstore CI signing lands (Theorem 5.3). The a11oy banned-token Doctrine gate is a known red CI item being tracked, disclosed in Chapter 13.*

Frontier claim. *Baking the router into the anatomy makes the model-selection decision auditable end-to-end — closing the gap where conventional routers hide* which *model decided behind an opaque dispatch.*

Chapter 12

Mesh Wiring: Wires A–H

The organs are interconnected by a mesh of eight wires (A–H). This chapter records each wire's endpoints and its *honest* status. Two wires are not yet implemented; we render them PENDING rather than claiming a complete mesh.

12.1 Wire inventory

Wire	Endpoints	Carries	Status
A	YAWAR internal	receipt bus (cord append + snapshot)	LIVE (internal)
B	a11oy ↔ sentra	immune verdicts / halt signals	**LIVE**
C	a11oy ↔ rosie	receipt stream to operator console	**LIVE**
D	W3C traceparent across mesh	distributed-trace correlation	**PENDING**
E	a11oy ↔ amaru cortex	doctrine/theorem enforcement	LIVE
F	a11oy ↔ vessels/killinchu khipu	telemetry → receipt cords	LIVE
G	brain-jack mesh (YUYAY ↔ UNAY)	cross-session memory relay	**PENDING**
H	lean-kernel backplane (OTel ↔ AMARU)	observability + kernel binding	LIVE

12.2 Per-wire notes

Wire A — receipt bus. The internal YAWAR bus that carries cord appends and frozen snapshots (Chapter 5). It is the substrate's spinal cord; everything that happens emits onto it.

Wire B — a11oy↔sentra (LIVE). Carries immune verdicts and halt signals over HTTP between the gate console and the inline immune layer (Chapter 6). Verified live.

Wire C — a11oy↔rosie (LIVE). Streams receipts to the operator console so the Body of Evidence is watchable in the 3D anatomy (Chapter 10). Verified live.

Wire D — W3C traceparent (PENDING). The intended distributed-trace correlation across the mesh via the W3C `traceparent` header is *not yet implemented*. This is disclosed as the headline mesh gap; it is on the v21 roadmap (Chapter 16).

Wire E — a11oy↔amaru cortex (LIVE). Carries Doctrine enforcement and theorem references between the gate and the cortex; the cross-component Doctrine invariant is PROVEN (`doctrine_cross_invariant`).

Wire F — a11oy↔vessels/killinchu khipu (LIVE). Carries maritime/drone telemetry into receipt cords (Chapter 9).

Wire G — brain-jack mesh (PENDING). The YUYAY↔UNAY cross-session memory relay is *not yet implemented*; UNAY has no dedicated module on the remote, so this wire renders PENDING.

Wire H — lean-kernel backplane (LIVE). Binds observability (OTel VSP) to the cortex (AMARU) and to the Lean-kernel reference layer; the exporter runtime is verified (19/19 runtime tests pass).

Remark 12.1 (Honest mesh status)**.** *Two of eight wires (D, G) are PENDING; the rest are live or internal. We deliberately do not claim a complete mesh. The 3D visualization (Chapter 10) renders D and G as PENDING toggles, so the honest status is visible in the picture. The mesh is the next ratchet of v21 (Chapter 16).*

Frontier claim. *A receipt-carrying organ mesh with per-wire honest status is the infrastructure layer beneath the application Spaces — the "Anatomy-as-Infrastructure" positioning of Chapter 14.*

Chapter 13

Honest Disclosure Ledger

This chapter is the single source of truth for the substrate's honest status. It states the canonical Lean numbers, the per-theorem status table, the A2/A4 axiom delta, the deposit-hygiene corrections, the SLSA level, the signing status, and the mesh gap. Nothing in this thesis claims more than what is recorded here.

13.1 Canonical Lean numbers

The Lean corpus (root `lutar-lean`, 99 `.lean` files, pinned SHA `c7c0ba1`) carries:

- **749 declarations** (canonical doctrine count);
- **14 unique axioms** (15 raw, 1 duplicate);
- **163 `sorry` obligations = 112 baseline + 51 Putnam**.

Remark 13.1 (Number drift, disclosed)**.** *Some surfaces carry stale figures: the `lutar-lean` README shows 752/15/160; the org `TRUST.md` historically showed 626/14/189. The canonical numbers are **749/14/163**, produced by `.github/scripts/lean_numbers.py` at the pinned SHA. The drift is being closed by wiring that script into a CI gate that fails on README/TRUST drift. The Putnam sorries total exactly 51 (P_A1=1, P_A2=8, P_A3=0, P_A4=4, P_A5=6, P_A6=2, P_B1–B6 = 6,6,4,6,6,2).*

13.2 The 14 canonical axioms

Axiom name	Role
`MomentSubGaussian`	PAC-Bayes concentration
`audit_reidemeister_invariance`	knot governance (R-moves)
`canonicalReceipt`	receipt canonicalization
`chromotopology_code_bijection`	QEC code algebra
`gleason_length_mod_8`	Gleason mod-8 (KS-18 witness)
`klDivergence_nonneg`	information-theoretic nonnegativity
`lambda_schur_concave_n_axis`	Schur-concavity of Λ (A4 corollary)
`lambda_stationary_unique`	uniqueness scaffold (Conjecture 1)
`liu_hui_pi_converges`	π series convergence
`pinsker`	Pinsker inequality
`r1_invariance`	Reidemeister R1
`r2_invariance`	Reidemeister R2

Axiom name	Role
`sha256`	hash model
`sha256_inj`	hash injectivity (receipt binding)

13.3 Per-theorem status table

Result	Status	Note
A2 homogeneity (`IsHomogeneous`)	[PROVEN]	degree-1 sc
A4 boundedness (`IsBounded`)	[PROVEN]	maps into
Λ monotonicity (TH_V18_01)	[PROVEN]	vector-bac
Λ product formula (TH_V18_12)	[PROVEN]	vector-bac
Khipu checksum invariant (TH11)	[PROVEN]	sorry-free,
Doctrine cross-invariant	[PROVEN]	Wire E en
Replay determinism (`yuyay_v3`)	[PROVEN]	purity + fi
Schur-concavity	[AXIOM]	`lambda_sc`
KS-18 contextuality witness	[AXIOM]	`gleason_l`
Reidemeister invariance	[AXIOM]	R1/R2 axi
Uniform-weight bridge	[SORRY]	`UnifiedLa`
Cauchy n-D step	[SORRY]	`Uniquenes`
PAC-Bayes / Madhava bound	[SORRY]	4 sorries
Two-witness convergence	[SORRY]	`TwoWitnes`
Λ uniqueness (Conjecture 1)	[CONJECTURE]	*not Theore*
UNAY cross-session memory	[CONJECTURE]	no remote
Pacha-Λ / Khipu–Bekenstein / Yachay–Khipu	[PROPOSED – Lean obligations pending]	Chapter 8

13.4 The A2/A4 axiom delta from v3

Axiom	v3 reading	v20 honest form
A2	zero-pinning ($\Lambda = 0$ iff some $x_i = 0$)	`IsHomogeneous` (degree-1 scaling); zero-pinning is a *consequence*, not the axiom
A4	page-curve concavity	`IsBounded` (maps into $[0, 1]$); Schur-concavity carried separately as `lambda_schur_concave_n_axis`

The v3 readings were over-strong/vacuous as stated; v20 restates them in the honest forms above, preserving every downstream property as an explicit tracked obligation rather than baking it into an axiom.

13.5 Deposit-hygiene corrections

- **v18 Zenodo deposit carried the v17 body.** The deposit `10.5281/zenodo.20434276` labelled v18 in fact contained the v17 manuscript; this is corrected via the GitHub release of the present v20 bundle and the new Zenodo deposit (Chapter 15).

- **v18 deposit "zero sorry/axiom" was false.** The honest figures are 749/14/163 (Section 13.1).
- **12 false `@lean_status:GREEN` bindings corrected.** v20 corrects twelve in-code bindings that labelled conjectural/sorry-tagged results as GREEN; the corrected statuses are those of Section 13.3.

13.6 SLSA level and signing status

Supply-chain provenance is honestly **SLSA Level 1**. A workflow in PR #235 labelled "SLSA Level 3" and was reverted; the honest `TRUST.md` value is L1, and the org coordination docs confirm the intended truth is "SLSA L1 (not L3)." **Sigstore/cosign CI signing is pending**, so receipts are labelled `PLACEHOLDER` throughout (Theorem 5.3). The summation tamper-evidence (Theorem 5.2) is independent of signing and holds today; signing adds non-repudiation on top.

13.7 Mesh gap

Wire D (W3C traceparent) and Wire G (brain-jack memory relay) are **not yet implemented** (Chapter 12); the mesh is rendered with honest PENDING status. Several shipped Spaces carry red CI items (a11oy banned-token gate; sentra container-build/hf-sync; vessels Tests/DCO), disclosed here and tracked.

Remark 13.2 (The honesty arc)**.** *This ledger is the culmination of an honesty arc: v12 introduced uniqueness as a Theorem; v14 downgraded it to a Conjecture on discovering the axiom structure; v19 stated the demotion in the abstract and retracted the σ-algebra claim ("measurable governance operator" $\to$ "append-only audit log + Khipu Merkle DAG"), honestly raising the sorry count* 163 $\to$ 168 *in that draft; v20 carries the corrected canonical 749/14/163 and the full per-theorem table above. The thesis line is* honest-by-construction.

Chapter 14

The Anatomy as Frontier Category

This chapter states the positioning thesis: the substrate is not "five agents" but *infrastructure* — a formally-verified governance anatomy on which the agents run as reference implementations. We argue why this is a category nobody else owns and locate it against the defense-autonomy frontier.

14.1 Anatomy-as-Infrastructure

The temptation is to lead with the agents (AMARU, SENTRA, ROSIE, vessels/killinchu, a11oy). That frame is crowded: everyone ships agents. The defensible frame is **Anatomy-as-Infrastructure**: the moat is the formally-verified governance gate (Λ proved/conjectured in Lean 4, emitting a DSSE-signed receipt onto a hash-linked Khipu Merkle DAG), and the five agents are the *evidence* that the substrate runs real workloads — exactly the proof an infrastructure buyer wants.

Remark 14.1 (Why Position 2)**.** *Of the twelve named organs, nine resolve to real shipped, tested code, two are PARTIAL, and one (KANCHAY brand-projection) is concept-only — so the substrate is real enough to lead with as infrastructure, provided the honesty fixes of Chapter 13 ship. The strongest single asset is the Lean-proved Λ gate + DSSE receipt chain + Khipu Merkle DAG, a category nobody else owns.*

Novelty. *Positioning the verified governance gate as the* infrastructure layer *(with agents as reference implementations) rather than as one feature of one agent — the "cognition-layer moat" analogue to a cluster-layer platform.*

14.2 The cluster-layer / cognition-layer analogy

The defense-autonomy platform incumbent (Defense Unicorns) owns the *cluster layer*: it signs the container *image* at admission (cosign), packages, and deploys. The substrate owns the *cognition layer*: it signs — and arithmetically locks — every *decision* the AI made inside the container at run time. These are complementary, not competing: "UDS signs the box the software came in; we sign every decision the AI made inside it" [8].

14.3 Warhacker positioning: the two SZL-unique problems

Against the public Warhacker problem catalog, the substrate fits *natively* on two problems, contributes on two more, and is honestly out of scope on two. We lead with the two it owns.

Problem A — the watchman who cannot be trusted. Every available AI-oversight tool is itself a learned, probabilistic, bypassable monitor; the defense memo lists deception-to-avoid-shutdown and guardrail-bypass as core risks [8]. The substrate's "allow/deny" is a kernel-checked Λ-gate verdict, not a guess — *the only watchman whose verdict is a verified mathematical object.* (Honestly: the headline uniqueness is Conjecture 1; the gate's *purity and replay determinism*, however, are PROVEN, Theorem 3.2, which is what the oversight claim actually rests on.)

Problem B — the non-refutable Body of Evidence. A Warhacker judge asked for evidence that can be evaluated "in minutes," producing "a non-refutable Body of Evidence," not "plans" [8]. Conventional observability yields piles of deletable, unsigned logs. The substrate's DSSE-signed Khipu Merkle DAG enforces a sum-of-sums invariant (TH11, Theorem 5.2): the root must equal the summed pendants or the receipt is mathematically invalid. The `tamper_test` demo flips one decision value and the math rejects the DAG in under a second — the math, not a policy, makes the evidence non-refutable.

Frontier claim. *Both headline problems are reinforced by current defense guidance: formally-verified oversight answers the deception/bypass risk, and the summation-checked receipt answers the non-repudiation/comprehensive-artifact requirement [8].*

14.4 Honest scope of the positioning

We do not claim to solve the orbit-visualization or medical-screening Warhacker problems; claiming them would dilute a sharp, defensible story. The substrate's contribution to the System-of-Record/ATO flow is the *evidence artifact* the record ingests, not the human risk-acceptance decision itself. This honest boundary is itself part of the moat: an oversight tool that over-claims is, by its own standard, failing the calibration axis.

Chapter 15

Citation Web

This chapter records the DOI lineage, the cited prior art, and the cited code, so that every claim in this thesis is traceable to a Lean file:line, a thesis chapter, a DOI, or a published source.

15.1 DOI lineage

Concept DOI. `10.5281/zenodo.19944926` (the version-agnostic concept record) resolves to the latest version, which this thesis updates to v20.

Version	Zenodo DOI	Note
Concept	10.5281/zenodo.19944926	version-agnostic; points to latest (v20)
v18	10.5281/zenodo.20434276	*carried v17 body by error* (Section 13.5)
v18 software	10.5281/zenodo.20434308	Lutar v18.0.0 software
v10	10.5281/zenodo.20053163	audit-closure operator $\rho = 1.000$
v19	*(unminted)*	drafted, package ready, not deposited
v20	*(this deposit)*	**the citable version**

Per-version DOI deltas are recorded in the version ledger (`173_VERSION_LEDGER_DELTAS.md`). After this deposit, the concept DOI `10.5281/zenodo.19944926` is updated to point to v20 as latest.

15.2 Cited prior art

The substrate cites a coherent canon spanning entropy bounds, concentration, codes, topology, learning theory, and quantum information: Bekenstein's entropy bound [3] with Casini's relative-entropy proof [6] and the Hayden–Wang zero-bit refinement [12]; Reidemeister's knot moves [24] and Witten's Jones-polynomial QFT [26]; McAllester's PAC-Bayes [17] with Catoni localization [7]; Pinsker's inequality [23] and Hoeffding's concentration [13]; the Mādhava–Kerala series for the two-witness bound [21]; Kochen–Specker contextuality [16]; the Wheeler "it from bit" frame [25]; the Brienza–Kung–Santos–Bobocel–Grossmann Situated Wise Reasoning Scale [5]; the Bravyi et al. bivariate-bicycle codes [4]; and Google's Willow below-threshold result, whose error-suppression factor is literally named $\Lambda = 2.14$ [10]. Cauchy/Aczél functional-equation theory [1] underwrites the open uniqueness step; the Bible-code refutation [27, 18] anchors the pre-registration discipline; the khipu ethnography [14, 19] grounds the Merkle-summation DAG; the defense-autonomy guidance [8] frames the two SZL-unique problems; and the design-pattern transfers cite Anthropic's constitution [2], DeepMind's AlphaProof [9], NVIDIA NeMo Guardrails [20], and the Pasterski–Strominger–Zhiboedov triangle [22].

15.3 Cited code

Repository	Role
`szl-holdings/lutar-lean`	Lean 4 kernel (749/14/163)
`szl-holdings/ouroboros`	runtime + 32-module self-test
`szl-holdings/ouroboros-thesis`	this paper substrate + history
`szl-holdings/platform`	monorepo (RAG, gates, doctrine-runtime)
`szl-holdings/a11oy`	governance console + LLM router
`szl-holdings/amaru`	cortex / chakra sidecar
`szl-holdings/sentra`	inline immune layer
`szl-holdings/vessels`	maritime intelligence
`szl-holdings/killinchu`	drone application (new)
`szl-holdings/rosie`	operator console + 3D anatomy
`szl-holdings/uds-mesh`	mesh schema library
`szl-holdings/vsp-otel`	observability exporter (19/19)

Remark 15.1 (Reproducibility)**.** *The vector-backed* Λ *core is reproducible against the 10 golden vectors in* `reference-vectors.json`*;* `OUROBOROS_RUN_ALL.py` *runs the 32-module self-test (live: 32/32). The Lean build is pinned to Mathlib v4.13.0 / Lean v4.13.0 at SHA* `c7c0ba1`*. Downstream conjectures are intentionally* not *vector-backed, consistent with their* ***[CONJECTURE]/[SORRY]/[P – Lean obligations pending]*** *status.*

Chapter 16

Conclusion and Future Work

16.1 What v20 consolidates

v20 is the culmination of the v14–v19 line: a single citable record of a formally-verified anatomical substrate for agentic AI. It fixes the canonical Λ definition (Chapter 2), the 13-axis conjunctive gate (Chapter 3), the HUKLLA/YAWAR/SENTRA organs (Chapters 4 to 6), the Maxwell $M{=}0$ wiring (Chapter 7), three Quechua-named frontier constructs (Chapter 8), the Killinchu drone application (Chapter 9), the 3D substrate visualization (Chapter 10), the seven-tier LLM router (Chapter 11), the mesh wiring (Chapter 12), and — above all — the honest disclosure ledger (Chapter 13) that makes every claim traceable and every gap explicit.

The central result stands and is honestly bounded: the substrate furnishes the "watchman that cannot itself go off script" (a kernel-verified gate verdict, PROVEN for purity and replay determinism, conjectural for uniqueness) and the "non-refutable Body of Evidence" (a summation-checked Khipu Merkle DAG, PROVEN sorry-free as TH11). These are the two problems the defense-autonomy community asked for and that no signing-the-box approach provides.

16.2 The v21 roadmap

v21 is defined by closing the gaps disclosed in Chapter 13:

1. **Sovereign loop wired end-to-end against the 13-axis gate.** Wire the full self-grading Ouroboros loop so that every step is gate-checked, receipted, and replayed against `yuyay_v3` end-to-end.

2. **Sigstore CI signing complete.** Replace every `PLACEHOLDER` signature with a real Sigstore/cosign attestation, retiring the HUKLLA T09 placeholder condition and lifting provenance toward an honest L2/L3 roadmap.

3. **Wire D implementation.** Implement W3C `traceparent` correlation across the mesh (Chapter 12), and stand up the Wire G memory relay once UNAY has a dedicated module.

4. **Discharge the open `sorry` obligations.** Run an AlphaProof-style Lean RL loop [9] against the CAUCHY_ND step (Theorem 2.5), the Reidemeister invariance obligations, and the PAC-Bayes/Madhava bounds — turning Conjecture 1 toward Theorem 1 honestly, only when the kernel actually closes the gap.

5. **Ship the frontier constructs from [PROPOSED – Lean obligations pending] to PROVEN.** Mechanize Pacha-Λ monotonicity/append-only, the Khipu–Bekenstein cap, and the Yachay–Khipu Reidemeister obligation.

16.3 The migration is the next ratchet

The substrate has reached the point where the work is no longer adding new math but *migrating* the live system to cite the single canonical definitions this thesis fixes: one Λ across all call sites, one Doctrine v11 gate, one canonical Lean number set, one honest provenance level. *The migration is the next ratchet.* Each turn of the ratchet — a call site unified, a `sorry` discharged, a `PLACEHOLDER` signed — moves the substrate from honestly-bounded to honestly-complete, without ever passing off aspiration as fact.

> *The loop closes on itself: a system that grades its own outputs against a pre-registered Doctrine, records the grade as a tamper-evident receipt, and halts when a survival invariant fires — the Ouroboros, made formal, made honest, and made citable.*

Acknowledgements and provenance

This work is authored by Stephen P. Lutar Jr. (SZL Holdings), ORCID 0009-0001-0110-4173, under the Concept DOI 10.5281/zenodo.19944926, licensed CC-BY-4.0. Where older drafts referred to a "Mythos" layer, the canonical name is *Hatun-Willay*. All numbers and statuses in this document are those of the honest disclosure ledger (Chapter 13); any live surface still claiming "proved uniqueness" or "SLSA L3" is out of date and superseded by this record.

Bibliography

[1] János Aczél. *Lectures on Functional Equations and Their Applications*. Academic Press, 1966. Theorem 5.1 (multivariate Cauchy equation), cited by the open uniqueness step.

[2] Anthropic. Claude's new constitution, 2026. URL: https://www.anthropic.com/news/claude-new-constitution.

[3] Jacob D. Bekenstein. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, 23(2):287–298, 1981. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.287, doi:10.1103/PhysRevD.23.287.

[4] Sergey Bravyi, Andrew W. Cross, Jay M. Gambetta, Dmitri Maslov, Patrick Rall, and Theodore J. Yoder. High-threshold and low-overhead fault-tolerant quantum memory. *Nature*, 627:778–782, 2024. URL: https://www.nature.com/articles/s41586-024-07107-7, doi:10.1038/s41586-024-07107-7.

[5] Justin P. Brienza, Franki Y. H. Kung, Henri C. Santos, D. Ramona Bobocel, and Igor Grossmann. Wisdom, bias, and balance: Toward a process-sensitive measurement of wisdom-related cognition. *Journal of Personality and Social Psychology*, 115(6):1093–1126, 2018. doi:10.1037/pspp0000171.

[6] Horacio Casini. Relative entropy and the bekenstein bound. *Classical and Quantum Gravity*, 25:205021, 2008. URL: https://arxiv.org/abs/0804.2182, doi:10.1088/0264-9381/25/20/205021.

[7] Olivier Catoni. *Statistical Learning Theory and Stochastic Optimization (Saint-Flour 2001)*, volume 1851 of *Lecture Notes in Mathematics*. Springer, 2004. doi:10.1007/b99352.

[8] Department of War (U.S.). Careful adoption of agentic ai services, 2026. Policy memo, 30 April 2026. URL: https://media.defense.gov/2026/Apr/30/2003922823/-1/-1/0/CAREFUL%20ADOPTION%20OF%20AGENTIC%20AI%20SERVICES_FINAL.PDF.

[9] Google DeepMind. Ai solves international mathematical olympiad problems at silver-medal level (alphaproof and alphageometry 2), 2024. URL: https://deepmind.google/blog/ai-solves-imo-problems-at-silver-medal-level/.

[10] Google Quantum AI, Rajeev Acharya, et al. Quantum error correction below the surface code threshold. *Nature*, 2024. URL: https://www.nature.com/articles/s41586-024-08449-y, doi:10.1038/s41586-024-08449-y.

[11] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1934.

[12] Patrick Hayden and Jonathan Wang. What exactly does the bekenstein bound? *Quantum*, 2025. URL: https://arxiv.org/abs/2309.07436.

[13] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830.

[14] Sabine Hyland. Writing with twisted cords: the inscriptive capacity of andean khipus. *Current Anthropology*, 58(3), 2017. URL: https://research-repository.st-andrews.ac.uk/handle/10023/12326, doi:10.1086/691682.

[15] Stephen P. Lutar Jr. Unifiedlambda proposal: reconciling three divergent definitions of the lutar invariant, 2026. szl-holdings/lutar-lean, Lutar/UnifiedLambda.lean (PROPOSAL); Concept DOI 10.5281/zenodo.19944926. URL: https://doi.org/10.5281/zenodo.19944926.

[16] Simon Kochen and Ernst P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17:59–87, 1967. doi:10.1512/iumj.1968.17.17004.

[17] David A. McAllester. Pac-bayesian model averaging. In *Proceedings of the 12th Annual Conference on Computational Learning Theory (COLT)*, pages 164–170, 1999. URL: https://dl.acm.org/doi/10.1145/307400.307435, doi:10.1145/307400.307435.

[18] Brendan McKay, Dror Bar-Natan, Maya Bar-Hillel, and Gil Kalai. Solving the bible code puzzle. *Statistical Science*, 14(2):150–173, 1999. URL: https://projecteuclid.org/journals/statistical-science/volume-14/issue-2/Solving-the-Bible-Code-Puzzle/10.1214/ss/1009212243.full.

[19] Manuel Medrano et al. How can data science contribute to understanding the khipu code? *Latin American Antiquity*, 2023. URL: https://www.cambridge.org/core/journals/latin-american-antiquity/article/how-can-data-science-contribute-to-understanding-the-khipu-code/E2D0E68A0515F3F9582A23C63693F4CE.

[20] NVIDIA. Nemo guardrails: programmable guardrails for llm applications, 2024. URL: https://github.com/NVIDIA/NeMo-Guardrails.

[21] Mādhava of Sangamagrama (Kerala school). Infinite series for π and arctangent (via secondary literature), 14th century. Used for the two-witness alternating-series error bound; see MacTutor history of mathematics. URL: https://mathshistory.st-andrews.ac.uk/Biographies/Madhava/.

[22] Sabrina Pasterski, Andrew Strominger, and Alexander Zhiboedov. The infrared triangle: soft theorems, asymptotic symmetries, and memory effects, 2014. URL: https://en.wikipedia.org/wiki/Pasterski%E2%80%93Strominger%E2%80%93Zhiboedov_triangle.

[23] Mark S. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden-Day, 1964.

[24] Kurt Reidemeister. Elementare begründung der knotentheorie. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 5:24–32, 1927. doi:10.1007/BF02952507.

[25] John A. Wheeler. Information, physics, quantum: the search for links. In Wojciech H. Zurek, editor, *Complexity, Entropy, and the Physics of Information*. Addison-Wesley, 1990.

[26] Edward Witten. Quantum field theory and the jones polynomial. *Communications in Mathematical Physics*, 121(3):351–399, 1989. doi:10.1007/BF01217730.

[27] Doron Witztum, Eliyahu Rips, and Yoav Rosenberg. Equidistant letter sequences in the book of genesis. *Statistical Science*, 9(3):429–438, 1994. URL: https://projecteuclid.org/journals/statistical-science/volume-9/issue-3/Equidistant-Letter-Sequences-in-the-Book-of-Genesis/10.1214/ss/1177010393.full.