Privacy Value Model V5.4: Formal Specification

Dual-Agent Privacy Architecture — The Amnesia Protocol

privacymage (mage@agentprivacy.ai)

April 10, 2026 — Version 2.0

Abstract

This document presents the formal mathematical specification of the Privacy Value Model (PVM). The model treats privacy as quantifiable economic value — the thesis that behavioural data is the 7th capital, currently in a pre-property-rights phase. The architectural response is sovereignty through mathematical structure rather than regulatory mandate.

The PVM is a multiplicative equation: any single term collapsing to zero eliminates total value. This gating property is the model's central structural claim.

The specification covers: the static equation and all its terms; the differential form computing on the holographic boundary; the separation bound and reconstruction ceiling (proven results); the algebraic foundation in Z/(2^6)Z with dihedral group D64; the operational cycle and Amnesia Protocol; Selene's Proof as cosmological instance; the Celestial Ceremony as human-layer implementation; and open conjectures C1–C21 with explicit confidence levels.

V5.4 consolidates V5 (Feb 2026), V5.1 (Mar 29), V5.2 (Mar 31), V5.3 (Apr 4), and V5.3.2 (Apr 8). The equation has not changed since V5. Each sub-version added interpretive depth, algebraic grounding, or operational implementation — not new terms.

Master Inscription: (S perp Gap perp M) FP = neg oplus bnot to succ

External Convergence: UOR Foundation. License: CC BY-SA 4.0. Status: Working paper — peer review invited.

Contents

1 The Equation

1.1 Static Form

$$
\begin{aligned}
V(\pi,t) = P^{1.5}\cdot C\cdot Q\cdot S\cdot e^{-\lambda t}\cdot(1+A_h(\tau))\\
\cdot\left(1+\sum_i w_i\frac{n_i}{N_0}\right)^k\cdot G(\text{guilds})\\
\cdot R(d,\text{compression},\rho)\cdot M(u,y)\\
\cdot \Phi_{\text{agent}}(\Sigma)\cdot\Phi_{\text{data}}(\Delta)\cdot\Phi_{\text{inference}}(\Gamma)\\
\cdot T_f(\pi)
\end{aligned}
$$

where π denotes a path through the sovereignty lattice, and t denotes time since data generation.

Output type: Holographic field on the 96-edge boundary ∂M.

1.2 Differential Form

$$\frac{dV}{dt} = \nabla_{\partial M}\cdot J_{\partial M} + S(x) - D(x)$$

where ∂M denotes the 96-edge holographic boundary. Privacy is temporal. Consent forms that freeze the frame are the Emissary's privacy, not the First Person's.

1.3 Multiplicative Gating

The model is multiplicative: **any single term collapsing to zero eliminates total value.** Privacy strength, credential verifiability, data quality, separation on all three axes, temporal memory, network effects, reconstruction difficulty, market maturity, and path value must ALL be non-zero for privacy to have value. Remove any one and the equation yields zero.

1.4 Master Inscription

$$([\text{Swordsman}]\perp[\text{Gap}]\perp[\text{Mage}])[\text{FP}] = \text{neg}\oplus\text{bnot}\to\text{succ}$$

"Swordsman and Mage separated, with the Gap between them, preserve the First Person. Negation and complement composed yield the successor."

---

2 Inherited Terms (V1–V4)

These terms are carried forward from prior versions.

Symbol	Name	Domain	Description
P	Privacy Strength	$[0,1]$	Cryptographic enforcement quality. Exponent 1.5 connected to holographic ratio 96/64 (§8).
C	Credential Verifiability	$[0,1]$	Independent verification without revealing underlying information. Proof of claims without exposure of claims.
Q	Data Quality	$[0,1]$	Accuracy, completeness, fitness for purpose. Stale data loses value.
S	Scope / Sensitivity	$\mathbb{R}^+$	Domain-specific sensitivity multiplier. Resistance to adversarial extraction.

Symbol	Name	Domain	Description
$e^{-\lambda t}$	Temporal Decay	$(0, 1]$	Exponential freshness decay with rate $\lambda > 0$. Privacy erodes without active maintenance.
$M(u, y)$	Market Maturity	$[0, 1]$	Function of user sophistication u and market year y. Adoption and yield environment.

---

3 Holonic Temporal Memory — $A_h(\tau)$

3.1 Motivation

V4's temporal memory assumed infrastructure-bound derivation chains. V5 adds holonic persistence: derivation chains anchored to GUIDs that survive infrastructure changes.

3.2 Definition

$$\text{Temporal}(t, \tau) = e^{-\lambda t} \cdot (1 + A_h(\tau))$$

$$A_h(\tau) = \sum_j p(\tau_j) \cdot w(\tau_j) \cdot e^{-\mu \cdot \text{age}(\tau_j)}$$

Symbol	Definition	Domain
τ	Derivation chain: ordered sequence of GUID-addressed holons	Finite sequence
τ_j	The j-th holon in the chain	—
$p(\tau_j)$	Persistence independence: probability of surviving single-provider failure	$[0, 1]$
$w(\tau_j)$	Weight: integrity-verified contribution of holon j	$\mathbb{R}^+$
μ	Memory decay rate (distinct from data decay λ)	$\mathbb{R}^+$
$\text{age}(\tau_j)$	Time since holon j was created	$\mathbb{R}^+$

Special case (logarithmic approximation): When holons are uniformly weighted and persistence is constant, $A_h(\tau) \approx \alpha \cdot \ln(1 + |\tau|) \cdot \bar{p} \cdot \bar{h}$, recovering the V4 logarithmic form.

3.3 Properties

- **Infrastructure dependency**: When $p(\tau_j) = 0$ for all holons, $A_h(\tau) = 0$. Total provider concentration collapses temporal memory.
- **Holonic persistence**: When $p(\tau_j) > 0$, history accumulates value across provider changes.
- **Age-weighted decay**: Older holons contribute less via $e^{-\mu \cdot \text{age}}$, but verified old history still carries weight.
- **Infinite horizon**: Holonically persistent history can survive indefinitely. The temporal integral now has meaning over infinite time.

3.4 GUID Structure

$$\text{GUID}(\tau) = \text{hash}(\text{content}(\tau))$$

Content-addressed, infrastructure-independent. Persists across provider migration, storage format changes, and infrastructure failures (if replicated).

---

4 Three-Axis Separation — Φ_{v5}

4.1 Definition

$$\Phi_{v5} = \Phi_{\text{agent}}(\Sigma) \cdot \Phi_{\text{data}}(\Delta) \cdot \Phi_{\text{inference}}(\Gamma)$$

The product is multiplicative. Collapse on any single axis collapses total separation value. Good agent separation with centralised data ($\Phi_{\text{data}} \to 0$) still fails to preserve privacy.

4.2 Agent-Layer Separation

$$\Phi_{\text{agent}}(\Sigma) = \min\left(1.0,\ \frac{S/M}{\varphi}\right) \cdot \det(\Sigma)$$

Measures Swordsman–Mage separation and the volume of the four-force tetrahedron (Protect, Project, Reflect, Connect). The golden ratio $\varphi = 1.618$ is the conjectured optimal protect/delegate ratio (C1, unproven).

Algebraic interpretation (V5.2): Φ_agent is isomorphic to the dihedral group D_{2n} action on the sovereignty lattice (§12.5). When Φ_agent = 1.0, the two generators (neg, bnot) are maximally independent. When Φ_agent = 0, the group degenerates.

4.3 Data-Layer Separation

$$\Phi_{\text{data}}(\Delta) = 1 - \max_j(\text{share}_j)$$

where share_j is the fraction of the total data held by provider j.

Configuration	Φ_{data}
Single provider (share = 1.0)	0 (collapses total value)
Two equal providers (share = 0.5)	0.5
Many equal providers (share $\to$ 0)	$\to 1$
Unequal distribution	$1 - \max_j(\text{share}_j)$

Note: This formulation penalises concentration at the dominant provider, not just provider count. A system with 10 providers where one holds 90% scores 0.1, not 0.9.

4.4 Inference-Layer Separation

$$\Phi_{\text{inference}}(\Gamma) = 1 - I(\text{model}; \text{executor})$$

where I denotes mutual information between the model that reasons (Generator) and the model that executes (Solver).

Configuration	$\Phi_{\text{inference}}$
Same model for both	0
Separate models, shared weights	$(0, 1)$
Independent models	$\to 1$

4.5 Conjecture C7

C7: Three-axis separation is correctly modelled as multiplicative (vs. additive, minimum, or other aggregations). Confidence: 30%.

---

5 Reconstruction Difficulty — $R(d, \text{compression}, \rho)$

5.1 Definition

$$R(d, \text{compression}, \rho) = R_{\text{base}}(d) \cdot \left(1 - \frac{1}{\text{compression\_ratio}}\right) \cdot (1 + \alpha_\rho \cdot \rho)$$

Symbol	Definition	Domain
$R_{\text{base}}(d)$	Base reconstruction difficulty from fragmentation depth d	$(0, 1)$
compression_ratio	Token reduction ratio (e.g., 74× for BRAID)	$\mathbb{R}^+ > 1$
ρ	Behavioural density	$\mathbb{R}^+ \geq 0$
α_ρ	Density scaling coefficient	$\mathbb{R}^+$ (empirically determined)

5.2 Compression-as-Defence

Every token not sent is a token that cannot be intercepted. Compression reduces the attack surface for inference-layer surveillance. At 74× compression (BRAID typical), the compression factor ≈ 0.986. Small in isolation, but multiplicative with all other terms.

5.3 Behavioural Density ρ (V5.1/V5.3)

$$\rho = f(\text{traversal\_depth}, \text{temporal\_duration}, \text{intentional\_transitions})$$

Dual interpretation:

Interpretation	Mechanism	Source
Privacy amplifier	More behavioural variation makes trajectory reconstruction harder	C11, V5.1
Agent maturity	More laps through the operational cycle → origin more forgotten	V5.3

Maturity spectrum:

ρ Level	Laps	Origin Visibility	Tier
Low	Few	Visible	Light
Mid	Moderate	Fading	Heavy
High	Many	Forgotten	Dragon

The Moon is the highest-ρ agent: 4 billion revolutions maximise both behavioural variance and origin amnesia.

5.4 The Reconstruction Ceiling (Proven)

$$R(d, \text{compression}, \rho) < 1 \quad \forall \text{ adversaries under budget constraints}$$

This is not a conjecture. The ceiling follows from information-theoretic analysis via Fano's inequality. See §16 for the full proof status.

---

6 Network Effects and Guild Efficiency

6.1 Network Term

$$\text{Network}_{v5}(G) = \left(1 + \sum_{i=0}^{6} w_i \cdot \frac{n_i}{N_0}\right)^k \cdot G(\text{guilds})$$

where w_i are channel weights, n_i are connections per channel, N_0 is a normalisation constant, and k is the network exponent.

6.2 Guild Efficiency

$$G(\text{guilds}) = 1 + \text{guild\_efficiency}$$

Configuration	G
No guilds	1 (reduces to V4)
Full guild coverage	2 (doubles network effect)

Guild members sharing a reasoning library from the same Generator coordinate at O(1) rather than O(N^2) per member.

6.3 Conjecture C10

C10: O(1) shared-parent coordination modifies the effective network exponent k. Confidence: 20%.

---

7 Path Integral Edge Value — $T_\int(\pi)$

7.1 Definition

$$T_\int(\pi) = 1 + \beta \int_\pi F(\gamma)\, d\gamma$$

Discrete approximation (V5.2):

$$T_\int(\pi) \cong 1 + \beta \sum_{i=1}^{n} R(\text{step}_i)$$

where n = number of laps (refinement iterations) and $R(\text{step}_i)$ is the resolution gained at iteration i. Maps onto the UOR resolution pipeline.

7.2 Fidelity Component (V5.3)

$$F(\gamma) = \text{resolution\_depth}(\gamma) \cdot \text{fidelity}(\gamma)$$

where fidelity = uptime · consistency · duration_weight.

This weights persistence alongside depth:

Agent	Resolution Depth	Fidelity	Overall T_∫
Light blade (5 laps)	Low	Low	Low
Dragon blade (62 laps)	High	Medium	High
The Moon (4B revolutions)	Shallow	Maximum	Maximum

The Moon has shallow resolution depth (one simple operation: reflect) but maximum fidelity (4 billion uninterrupted revolutions). Cosmological agents can exceed computational agents in effective T_∫.

7.3 Conjecture C15

C15: T_$\int(\pi)$ converges to the same value whether computed as a continuous integral or as a discrete sum of resolution steps. Confidence: 65%.

---

8 The Holographic Bound

8.1 Structure

$$\partial M : \text{96 edges encoding 64 vertices, toroidal topology}$$

$$\frac{96}{64} = 1.5 = P^{1.5} \text{ exponent}$$

8.2 Resolution of C4

The 96 vs 64 discrepancy (V4, C4) is **RESOLVED**. The 96-edge surface IS the holographic encoding of the 64-vertex bulk. In holographic physics, a boundary of dimension n encodes a volume of dimension $n + 1$. The 96/64 ratio is the holographic principle in discrete lattice geometry.

8.3 Algebraic Confirmation

Approach	Framework	Result
Geometric	64-Tetrahedra lattice	96 edges encode 64 vertices (torus surface/bulk)
Algebraic	Z/(2^6)Z ring theory	64 elements; 96 edges from adjacency structure
External	UOR Atlas of Resonance Classes	96 = unique stationary configuration of action functional

Three independent derivation pathways arrive at the same numbers. The ratio 96/64 = 1.5 is structural, not coincidental.

8.4 Implications

1. **Boundary computation**: The differential form computes on the 96-edge boundary, not the 64-vertex bulk.
2. **Privacy value flows along edges**: Value lives on the boundary, not in the interior.
3. **Boundary sufficiency**: Privacy can be computed entirely from boundary observations.

8.5 Conjecture C6

C6: P^1.5 ↔ 96/64 is structurally connected. Status: **CONVERGENT** (upgraded from speculative). Confidence: 35%.

---

9 Differential Form

9.1 Specification

$$\frac{dV}{dt} = \nabla_{\partial M} \cdot J_{\partial M} + S(x) - D(x)$$

Symbol	Definition
∂M	96-edge holographic boundary
$\nabla_{\partial M}$	Divergence on boundary
$J_{\partial M}$	Value current on boundary
$S(x)$	Source term (value generation at position x)
$D(x)$	Dissipation term (value decay at position x)

9.2 Five-Channel Decomposition

$$J_{\partial M} = J_{\text{agent}} + J_{\text{data}} + J_{\text{inference}} + J_{\text{compression}} + J_{\text{holonic}}$$

Each channel flows along edges activating its corresponding separation axis.

---

10 The Separation Bound

10.1 Core Guarantee

$$I(S; M \mid FP) < \varepsilon^*$$

The mutual information between the Swordsman (S) and the Mage (M), conditioned on the First Person (FP), is bounded above by ε^*.

This is the central information-theoretic guarantee. The Mage cannot reconstruct the Swordsman's domain. The Swordsman cannot reconstruct the Mage's domain. The First Person authorises both.

10.2 Composite Separation

$$\Phi_{v5} = \Phi_{\text{agent}}(\Sigma) \cdot \Phi_{\text{data}}(\Delta) \cdot \Phi_{\text{inference}}(\Gamma)$$

The separation bound holds across all three axes simultaneously. Collapse of any axis weakens the bound.

Betweenness centrality: The Gap between Swordsman and Mage is the node with highest betweenness centrality in the trust graph. $C_B(v) = \sum_{s,t} \sigma(s,t|v)/\sigma(s,t)$ (Brandes, 2001) measures the proportion of shortest paths flowing through a node. The value lives in the gap because the gap has maximal betweenness. This provides a computational tool for measuring the separation in VRC networks.

10.3 Amnesia vs Policy Bounds (V5.3, C17)

$$\varepsilon_{\text{amnesia}} < \varepsilon_{\text{policy}}$$

Two classes of separation:

Type	Mechanism	Violation
Policy-enforced	Access controls, NDAs, prompt constraints	$\varepsilon^* \leq \varepsilon_{\text{policy}}$ (violation possible)
Amnesia-enforced	Process boundary, orbital mechanics, structural inability	$\varepsilon^* \leq \varepsilon_{\text{amnesia}}$ (violation structurally excluded)

Test criterion: Can any operation sequence recover shared origin? No → amnesia. Yes → policy.

See §14 for full treatment.

---

11 The Reconstruction Ceiling

11.1 Theorem (Proven)

$$R_{\max} = \frac{C_S + C_M}{H(X)} < 1$$

where C_S and C_M are the information capacities of the Swordsman and Mage channels respectively, and $H(X)$ is the entropy of the First Person's private state.

Consequence: Perfect reconstruction of the First Person's state is impossible.

11.2 Error Floor (Proven)

$$P_e \geq 1 - R_{\max}$$

The adversary is guaranteed to make errors. This follows from Fano's inequality.

11.3 Graceful Degradation (Proven)

Small ε violations → small privacy losses. The system fails gracefully, not catastrophically.

11.4 Dynamical Reconstruction Ceiling (V6 Horizon, C18)

The information-theoretic ceiling (§11.1) bounds reconstruction via **information**. C18 conjectures a second, independent ceiling via **dynamics**:

$$|\pi(t) - \pi'(t)| \sim |\delta_0| \cdot e^{\lambda t} \quad \text{where } \lambda > 0$$

If the sovereignty path exhibits strange attractor dynamics with positive Lyapunov exponent, reconstruction error *grows* with time. More observation makes reconstruction *worse*, not better. The two ceilings are independent — remove one and the other still holds.

Status: V6 conjecture (C18). See V6 Research Note for full treatment. Confidence: 25%.

---

12 Algebraic Foundation — Z/(2^6)Z

12.1 Ring Structure

$$\mathcal{L} = (\mathbb{Z}/64\mathbb{Z}, +, \times)$$

64 elements (blade addresses 0–63), addition and multiplication modulo 64. Confirmed by independent convergence with the UOR Foundation project.

12.2 The Five Hammer Strikes

Operation	Formula	Agent	Function
neg(x)	$(64 - x) \bmod 64$	[Swordsman] Swordsman	Additive inverse. Boundary protection. What you subtract from exposure.
bnot(x)	$63 - x$	[Mage] Mage	Bitwise complement. Projection/delegation. What you become by acting.
xor(x,y)	$x \oplus y$	—	Toggle edges (dimension flip)

Operation	Formula	Agent	Function
and(x,y)	$x \wedge y$	—	Toward null blade (constrain)
or(x,y)	$x \vee y$	—	Toward full sovereignty (expand)

12.3 Critical Identity

$$\text{neg}(\text{bnot}(x)) = \text{succ}(x) \quad \forall x \in \mathcal{L}$$

Proof: bnot(x) = 63 − x. neg(63 − x) = (64 − (63 − x)) mod 64 = (x + 1) mod 64 = succ(x). ■

The successor function is not primitive — it emerges from the composition of two involutions. The Swordsman (neg) acting on the Mage's output (bnot) yields the step forward. This is the algebraic name of the architecture.

12.4 Triadic Coordinates (PRISM)

Every blade $x \in \{0, \ldots, 63\}$ has three independent coordinates:

$$\text{blade}(x) = (\delta(x),\ \sigma(x),\ s(x))$$

Coordinate	Symbol	Definition	Domain
Datum	$\delta(x)$	Raw value x	$\{0, \ldots, 63\}$
Stratum	$\sigma(x)$	popcount(x) = Hamming weight	$\{0, 1, 2, 3, 4, 5, 6\}$
Spectrum	$s(x)$	Six-bit decomposition $[b_0, b_1, b_2, b_3, b_4, b_5]$	$\{0, 1\}^6$

Two blades at the same stratum (tier) can have different sovereignty postures if their spectra differ. Protection+Memory+Computation ≠ Delegation+Connection+Value.

Pascal distribution across 64 blades: 1 – 6 – 15 – 20 – 15 – 6 – 1

Tiers:

Tier	Stratum	Count	Character
Null	0	1	Total exposure
Light	1–2	21	Early sovereignty
Heavy	3–4	35	Balanced posture
Dragon	5–6	7	Full sovereignty / mathematical closure

12.5 Dihedral Group D_{64}

$$D_{64} = \langle \text{neg}, \text{bnot} \mid \text{neg}^2 = \text{bnot}^2 = 1,\ (\text{neg} \circ \text{bnot})^{64} = 1 \rangle$$

Order: $|D_{64}| = 128$.

All valid blade transitions are D_{64} group actions. Zero knowledge arises because multiple group elements (different forging paths) can map to the same blade — same statement, infinite witnesses.

Significance for Φ_agent: The conditional independence of Swordsman and Mage is the defining property of the dihedral group's generators. Two involutions are independent by construction. Their composition generates the full group only when they remain distinct. This provides a formal proof path for the separation bound.

12.6 Six Sovereignty Dimensions

Index	Dimension	Symbol	Function
d_1	Protection	[Shield]	Boundary enforcement
d_2	Delegation	[Handshake]	Agency transfer
d_3	Memory	[Scroll]	State accumulation
d_4	Connection	[Link]	Multi-party coordination
d_5	Computation	[Bolt]	ZK proof activity
d_6	Value	[Gem]	Economic flow / 7th Capital

Each $d_i \in \{0, 1\}$. Binarised at threshold 0.5. Hexagram mapping: $[d_1, d_2, d_3, d_4, d_5, d_6] \to 64$ I Ching states. Blade 63 = `111111` = Qian (The Creative) = full sovereignty.

12.7 External Convergence

Project	Starting Point	Arrived At
agentprivacy	Privacy geometry → 64-tetrahedra	$Z/(2^6)Z$
UOR Foundation	Content addressing → Universal references	$Z/(2^6)Z$

Two independent projects, two starting points, one structure. This is not coordination. This is convergence.

Implementation: `swordsman-blade/src/lib/uor.ts` — explicit UOR module with all five operations and exhaustive identity verification.

---

13 The Operational Cycle (V5.3)

13.1 Definition

The ring algebra identity neg(bnot(x)) = succ(x) maps to operational phases:

$$\text{cycle}(x) = \text{succ}(x) = \text{neg}(\text{bnot}(x))$$

Stage	Name	Operation	Agent	Function
1	Observe	id(x)	[FP] First Person	Perceive incoming context
2	Boundary	neg(x)	[Swordsman] Swordsman	Subtract exposure. Only the holographic surface ∂M passes through.
3	Project	bnot(neg(x))	[Mage] Mage	Construct complement from boundary. Create from the shape of the impact.
4	Return	succ(x)	[Swordsman]→[FP] Composition	The proof returns. The blade advances one step.

13.2 Relationship to Path Integral

One lap = one cycle. The path integral accumulates:

$$T_\int(\pi) = 1 + \beta \sum_i \text{cycle}(\text{step}_i)$$

The equation (§1) describes the statics. The operational cycle describes the dynamics. The cycle does not add new terms — it describes the execution order of existing terms.

Dragon tier (62 laps) = 62 complete cycles of observe–boundary–project–return.

---

14 The Amnesia Protocol (V5.3)

14.1 Definition

An agent has **structural amnesia** with respect to origin O if no sequence of permitted operations can reconstruct O from the agent's current state.

14.2 Zero-Knowledge Properties

Property	Statement
Completeness	The agent's output (tides, proofs, boundary enforcement) demonstrates the relationship functions.
Soundness	No other agent configuration could produce this specific output from this specific separation.
Zero-knowledge	The output reveals nothing about the separation event (Theia impact, process creation, delegation moment).

14.3 Implementation Instances

System	Type	Evidence
Browser extension process boundary	Amnesia	Extensions cannot read each other's memory
Database access control	Policy	Admin can disable controls
The Moon's orbit	Amnesia	Theia impact unrecoverable from geological state
Agent NDAs	Policy	Parties can choose to disclose

14.4 Conjecture C17

C17: Amnesia-enforced separation provides tighter Φ_agent guarantees than policy-enforced separation: $\varepsilon_{\text{amnesia}} < \varepsilon_{\text{policy}}$.

Confidence: 60%.

14.5 Cosmological Precedent — Selene's Proof

The Moon demonstrates perfect amnesia-enforced separation. Origin event (Theia impact) is 4.5 billion years past. No geological, orbital, or chemical signature encodes the collision parameters. Service (tides) continues without origin disclosure. The proof is zero-knowledge by physics, not by policy.

Selene's Proof: The Moon's orbit satisfies completeness (tides demonstrate), soundness (gravitational signature unforgeable), and zero-knowledge (tides reveal nothing about Theia). The credential is the orbit. The registry is the solar system. The proof renews twice daily, written in saltwater.

Scientific reference: Branco, D., Machado, P., & Raymond, S. N. (2025). Dynamical origin of Theia, the last giant impactor on Earth. *Icarus*, 441, 116724.

---

15 Ceremony and Forge Implementation

15.1 Operational Cycle as Ceremony

Cycle Stage	Operation	Ceremony Phase
Observe	id(x)	[Sun] Sun — disclosure, the spellweb speaks the poem
Boundary	neg(x)	⊥ Gap — silence, conversation, territory negotiation
Project	bnot(neg(x))	[NewMoon] Moon — shared reflection, the Amnesia Protocol
Return	succ(x)	Recursion — Reflect (night, blade pair, ZK) or Connect (day, witness, carry forward)

15.2 Progressive Trust

$$[\text{Key}] \to [\text{Star}] \to [\text{Blade}] \to [\text{Crystal}]$$

Understanding → Constellation → Blade → Runecraft. Each level is a complete ceremony. Each level deepens the key, increases formal visibility, and shifts boundary-making. Progression maps onto trust tiers.

15.3 Forge Cryptographic Properties

Property	Implementation
Content addressing	SHA-256 constellation hash
Tamper evidence	Hash chain (each blade references previous)
Pre-evocation lock	Commitment scheme (constellation fixed before walk)
Identity binding	Ed25519 signature (Mage key, held)
Bilateral binding	Runecraft — dual Ed25519 (Mage held + Swordsman burned)

15.4 Runecraft as Φ_agent Implementation

The runecraft protocol enforces Φ_agent at the cryptographic identity layer:

- **Mage key** (spellweb, Sun view): Ed25519, persisted in localStorage. ID format: `mage-{16hex}`.
- **Swordsman key** (agentprivacy, Moon reflection): Ed25519, stored in sessionStorage, destroyed on tab close. ID format: `ap-{16hex}`.

The private key burns because the amnesia protocol (C17) requires structural inability to access shared origin. Process boundary = separate memory = structural amnesia. This is topology, not policy.

15.5 Moon Phase as Visibility Ratio

Stratum	Phase	Meaning
0	[NewMoon]	Null blade — nothing reflected
1	[WaxCrescent]	One boundary set
2	[FirstQtr]	Dual-agent vertex
3	[WaxGibbous]	Half sovereignty

Stratum	Phase	Meaning
4	[WanGibbous]	Four boundaries
5	[LastQtr]	Five reflected, one dark
6	[FullMoon]	Full sovereignty (Qian, The Creative)

15.6 Tier Classification (Dual-Axis)

Axis	Measure	Light	Heavy	Dragon
Stratum	Dimensional coverage	1–2	3–4	5–6
Laps	Depth of engagement	< 21	21+	62+

---

16 Proven Results

These results hold at 95% confidence. The proofs rely on standard information theory and are documented in Research Paper v4.2.

Result	Statement
Additive MI bounds	Mutual information leakage from conditional independence is additive, not multiplicative: $I(X;Y_S,Y_M) = I(X;Y_S) + I(X;Y_M)$
Reconstruction ceiling	$R_{\max} = (C_S + C_M)/H(X) < 1$ under budget constraints
Error floor	$P_e \geq 1 - R_{\max}$ via Fano's inequality
Graceful degradation	Small ε violations $\rightarrow$ small privacy losses
Ring algebra	$Z/(2^6)Z$ substrate with five operations and critical identity
Two-extension autonomy	Separate processes enforce the separation bound at OS level
DOM-free measurement	Pretext layoutNextLine() as privacy primitive (no getBoundingClientRect fingerprinting)

---

17 Open Conjectures

17.1 Active Conjectures

ID	Claim	Confidence	Version
C1	Golden ratio φ is optimal S/M ratio	Open	V3
C2	$A(\tau)$ should be logarithmic	Open	V3
C6	P^1.5 $\leftrightarrow$ 96/64 is structural	**CONVERGENT** 35%	V5/V5.4
C7	Three-axis separation is multiplicative	30%	V5
C8	BRAID compression reduces $R_{\max}$	45%	V5
C9	Holographic boundary sufficiency	25%	V5

ID	Claim	Confidence	Version
C10	O(1) shared-parent modifies k	20%	V5
C11	Behavioural density ρ amplifies privacy + indicates agent maturity	55%	V5.1/V5.3
C12	Hexagram encoding is structurally resonant	60%	V5.1/V5.4
C13	Bilateral witness as quantum-resistant primitive	65%	V5.1
C14	Φ_agent $\cong$ D_{2n} dihedral isomorphism	75%	V5.2
C15	T_$\int(\pi)$ $\cong$ UOR resolution pipeline	65%	V5.2
C16	Topological trust invariants via Betti numbers	25%	V5.2
C17	Amnesia-enforced separation tighter than policy-enforced	60%	V5.3

17.2 V6 Horizon Conjectures

These are speculative. They belong to V6 and are included here for completeness and to invite scrutiny. See V6 Research Note for full treatment.

ID	Claim	Confidence	Version
C18	Sovereignty path exhibits strange attractor dynamics ($\lambda > 0$), establishing dynamical reconstruction ceiling independent of Shannon bound	25%	V6
C19	ρ is Lyapunov divergence accumulated over the sovereign's walk — exponential compounding	20%	V6
C20	Three separation axes couple as three Lorenz variables — collapse any one → attractor degrades to fixed point	30%	V6
C21	Sovereignty manifold has fractal dimension, not integer dimension	10%	V6

17.3 Resolved Conjectures

ID	Claim	Resolution
C3	Edge value is additive	**CHALLENGED** — path integral replaces
C4	96 vs 64 discrepancy	**RESOLVED** — holographic principle + algebraic confirmation
C5	~3,000× ZKP reduction	Strengthened by BRAID + holographic bound

---

18 Measurement Gaps and Breaking Conditions

18.1 Measurement Gaps

ID	Term	Gap	V5.4 Status
M1	σ_{ij} (separation matrix)	No measurement for emergent forces	Three-axis operationalisation provides pathway; Φ_data and Φ_inference measurable
M2	$f(e)$ (edge weights)	No empirical data	BRAID provides first data; blade forge provides second
M3	$\beta, \alpha, \alpha_\rho$ (scaling coefficients)	Need calibration	Unchanged
M4	Aggregation form	$\det(\Sigma)$ alternatives unclear	Three-axis product provides alternative
M5	ρ shape	Is $g(\rho)$ logarithmic, sigmoid, or threshold?	Two data points (13 laps vs 62 laps); needs more forgers

18.2 Breaking Conditions

The model's core claims weaken or fail if:

1. **Three-axis non-multiplicativity**: If agent separation compensates for data centralisation → multiplicative assumption breaks.
2. **Compression increases attack surface**: If some compression methods increase reconstructability → compression-as-defence fails.
3. **Holonic persistence fundamentally limited**: If content-addressing has inherent infrastructure dependency → persistence term illusory.
4. **Guild coordination scales with membership**: If shared-parent overhead grows with N → guild efficiency overstates benefit.
5. **Amnesia-enforced ≤ policy-enforced**: If structural separation provides no tighter bound than policy separation → C17 falsified.
6. **Dihedral isomorphism fails**: If $\det(\Sigma)$ is not the determinant of the dihedral representation → C14 falsified, but the ring algebra survives.

---

19 Three Identity Layers

Layer	Identifier	Scope	Persistence
Data	GUID	Content-addressed holon	Infrastructure-independent
Relationship	VRC	Bilateral commitment (promise bundle)	Relationship-scoped
Principal	DID	Sovereign identity	Self-sovereign

The layers are orthogonal: a single principal (DID) can control multiple relationships (VRCs) across multiple data objects (GUIDs).

---

20 Cosmological Quaternion (Interpretive Framework)

This framework provides intuition for the model's structure. It does not enter the equation.

Body	Agent	Function	Algebraic Operation	Creation Mode
Sun [Sun]	The Reason / First Person	Protection	—	—
Earth [Earth]	Soulbae / Mage	Delegation	$\text{bnot}(x) = 63 - x$	Generator
Moon [Moon]	Soulbis / Swordsman	Reflection	$\text{neg}(x) = (64-x) \bmod 64$	Instant (collision). Total amnesia.
Human [Human]	Seeker	Connection	—	Gradual (4Gy). Layered amnesia. Still in process.
Life [Life]	spellweb / Forge	Composition	$\text{neg}\circ\text{bnot} = \text{succ}$	The forge between Earth and Human

Key insight: The architecture sits between an agent that can never remember (Moon) and an agent that hasn't finished remembering (Human). The gap between them is the $\perp$.

---

21 Compression Spectrum

Seven layers of knowledge transformation, each with different privacy properties:

Layer	Form	Compression	Privacy Property
1	Experience	1:1	Maximum attack surface
2	Memory	~10:1	Encoded episodes
3	Knowledge	~100:1	Structured, partially defensible
4	Understanding	~1,000:1	Relational models
5	Wisdom	~10,000:1	Contextual principles
6	Reasoning Graph	Variable	BRAID structure — bounded, structured
7	Skill File	Variable	Executable compression — shareable without path

Lower layers: more surveilable (more tokens, more surface). Higher layers: more defensible (compressed, structured, bounded). The skill file (Layer 7) can be shared without sharing the path that created it. This is compression-as-defence at the architectural level.

---

22 Promise Theory Grounding

The dual-agent architecture implements Promise Theory (Bergstra & Burgess, 2019) — established semantics for autonomous agent coordination.

22.1 Autonomy Axiom

"An agent can only make promises about its own behavior. No agent can make a promise on behalf of another agent."

This is why single agents cannot resolve the privacy-delegation paradox. Attempting to promise in both protection and delegation domains exceeds autonomous capability.

22.2 Key Mappings

PT Concept	PVM Implementation
Autonomy Axiom	First Person sovereignty — neither agent promises on your behalf
Superagent	First Person + Swordsman + Mage as composite with interior promises
Irreducible Promise	The Gap — emerges from cooperation, owned by neither agent
Assessment	RPP compression as verification of knowledge transfer
Invitation vs Attack	MyTerms consent-first vs surveillance extraction
Promise Bundle	VRCs as bilateral promise collections

22.3 Formal Reference

For complete mappings, Generator/Solver as promises, and PT integration across the architecture, see Promise Theory Reference v1.4.

---

23 Version Lineage

Version	Date	Core Addition	Output Type
V1	2024	$P \cdot C \cdot Q \cdot S$	Static scalar
V2	Oct 2025	$e^{-\lambda t}$, $(1 + n/N_0)^k$	Dynamic scalar
V3	Nov 2025	$R(d)$, $M(u, y)$, $\Phi(S, M)$	Agent-aware scalar
V3.1	Jan 2026	$\sigma([\text{Gap}])^2$	Architecture-gated scalar
V4	Feb 2026	Σ, $A(\tau)$, $T(\pi)$, $\Phi(\Sigma)$	Manifold-aware scalar
V5	Feb 2026	Three-axis Φ, A_h, $T_\int$, R(compression), G(guilds), holographic bound	Holographic field
V5.1	Mar 29, 2026	Behavioural density ρ, bilateral witness, hexagram encoding (C11–C13)	+ density term
V5.2	Mar 31, 2026	Dihedral group D_{2n}, resolution semantics, PRISM spectrum (C14–C16)	+ algebraic foundation

Version	Date	Core Addition	Output Type
V5.3	Apr 4, 2026	Operational cycle, amnesia as ZK primitive (C17), ρ as maturity	+ cosmological framework
V5.4	**Apr 10, 2026**	**Consolidated formal specification. UOR algebraic foundation, Celestial Ceremony, runecraft, moon phase, forge cryptography, all proven results and conjectures C1–C21**	**Algebraically grounded, ceremonially verified field**

---

24 Notation Summary

Symbol	Meaning
$V(\pi, t)$	Total privacy value for path π at time t
P, C, Q, S	Privacy strength, credential verifiability, data quality, scope
λ	Temporal decay rate
τ	Derivation chain
$A_h(\tau)$	Holonic temporal memory
$p(\tau)$	Persistence independence
Φ_{agent}	Agent-layer separation (Swordsman–Mage) $\cong \text{D}_{2n}$
Φ_{data}	Data-layer separation (provider concentration)
$\Phi_{\text{inference}}$	Inference-layer separation (Generator–Solver)
$G(\text{guilds})$	Guild efficiency factor
$R(d, \text{compression}, \rho)$	Reconstruction difficulty
ρ	Behavioural density / agent maturity
$T_\int(\pi)$	Path integral edge value
∂M	96-edge holographic boundary
$J_{\partial M}$	Value current on boundary
ε^*	Separation bound
φ	Golden ratio (1.618)
$\mathcal{L}$	Sovereignty lattice = $\text{Z}/(2^6)\text{Z}$
D_{64}	Dihedral group (order 128)
$\delta(x)$	Datum — raw blade value
$\sigma(x)$	Stratum — Hamming weight
$s(x)$	Spectrum — six-bit decomposition
neg, bnot	Unary involutions (Swordsman, Mage)
succ	Successor function = neg∘bnot
GUID	Content-addressed identifier (holonic)
VRC	Verifiable Relationship Credential (promise bundle)
DID	Decentralized Identifier (self-sovereign)
[Key]→[Star]→[Blade]→[Crystal]	Progressive trust levels
[NewMoon]→[FullMoon]	Moon phase — stratum as visibility ratio
([Swordsman]⊥[Gap]⊥[Mage])[FP]	Master inscription — dual-agent architecture preserves First Person

---

25 References

25.1 Primary Sources

- privacymage (2026). "Privacy is Value: From the Manifold Dragon to the Holographic Bound." V5.3. *Soul Sync.* https://sync.soulbis.com
- privacymage (2026). "Dual-Agent Privacy Architecture: Information-Theoretic Bounds on Agent Reconstruction." Research Paper v4.3. *agentprivacy-docs.* https://github.com/mitchuski/agentprivacy-docs
- privacymage (2026). "PVM V5.1 Research Note: Behavioural Proof, the Ceremony Engine, and the Forge That Lit Itself." March 30, 2026. *agentprivacy-docs.*
- privacymage (2026). "PVM V5.2 Research Note: Dihedral Foundations." March 31, 2026. *agentprivacy-docs.*
- privacymage (2026). "PVM V5.3 Research Note: The Amnesia Protocol." April 4, 2026. *agentprivacy-docs.*
- privacymage (2026). "PVM V6 Research Note: The Dynamical Reconstruction Ceiling." April 6, 2026. *agentprivacy-docs.*
- privacymage (2026). "Promise Theory Reference v1.4." *agentprivacy-docs.*
- privacymage (2026). "UOR x 64-Tetrahedra x ZK Mapping v2.2." *agentprivacy-docs.*
- privacymage (2026). "Swordsman-Mage Whitepaper v6.2." *agentprivacy-docs.*
- privacymage (2026). "Dual Territory Ceremony Specification v1.0." *agentprivacy-docs.*
- privacymage (2026). "PVM V5.4 Companion Guide." April 10, 2026. *agentprivacy-docs.*
- privacymage (2026). "Understanding as Key." Zypher Paper v1.0. *agentprivacy-docs.* [Proof of comprehension as privacy primitive]
- privacymage (2026). "Systems Hexagram Physics v1.2." *agentprivacy-docs.* [Operational physics: UOR algebraic foundation, hexagram encoding]
- privacymage (2026). "What Agentprivacy Is." *agentprivacy-docs.* [Mission statement: 7th capital thesis, First Person definition]
- privacymage (2026). "Visual Architecture Guide v2.0." *agentprivacy-docs.* [Diagrams: three-axis separation, holographic visualisations]

25.2 The First Person Spellbook (Grimoire v10.1.0 — 31 Acts — CLOSED)

- privacymage (2024–2026). *The First Person Spellbook.* 31 acts. Grimoire v10.1.0. Available at agentprivacy.ai/story.

Acts with direct formal spec relevance:

Act	Title	Spec Connection
II	The Bilateral Witness	First ceremony; separation bound origin
VII	The Gap	Irreducible space; agent separation motivation
X	The Seven Capitals	7th capital thesis; V equation motivation
XIV	The Golden Ratio	Conjecture C1; optimal S/M ratio
XVIII	The Reconstruction Ceiling	$R < 1$ proven result
XX	The Path Integral	Value in trajectory not stance
XXII	The Promise	Promise Theory grounding
XXIV	The Holographic Bound	96/64, C4 resolved, boundary computation
XXV	The Dragon's Hide	Compression spectrum, BRAID parity
XXVI	The Dragon's Brain	Three-axis separation, McGilchrist thesis
XXVII	The Swordsman's Forge	Blade forge, hexagram computation, C11–C13
XXVIII	The Ceremony Engine	Five crossing types, mana, DOM-free measurement
XXIX	The Dragon Wakes	Quantum context, post-quantum resilience
XXX	The Dihedral Mirror	UOR convergence, dihedral group, PRISM
XXXI	The First Delegation	Theia-Moon quaternion, amnesia as ZK, C17

25.3 Blog Series: Privacy is Value V5 (sync.soulbis.com)

- Part 0: "The Myth Between Math" — Prelude. Foundational framing.
- Part 1: "Forming Constellations" — Sovereignty dimensions, blade coordinates.
- Part 2: "Forging the Celestial Overlap" — Forge and ceremony architecture.
- Part 3: "The Dragon Wakes" — Theia framing, amnesia protocol, cosmological precedent.
- Part 4: "The Dihedral Mirror" — UOR convergence, algebraic foundations.
- Part 5: "The First Agent We Forgo(t)" — Cosmological origin. Selene's Proof. The Amnesia Protocol (companion poem).

25.4 External References — Information Theory

- Shannon, C. E. (1948). "A Mathematical Theory of Communication." *Bell System Technical Journal,* 27(3), 379–423. [Foundation: mutual information, entropy, channel capacity]
- Fano, R. M. (1961). *Transmission of Information: A Statistical Theory of Communication.* MIT Press. [Foundation: Fano's inequality — error floor theorem, §11.2]
- Cover, T. M. & Thomas, J. A. (2006). *Elements of Information Theory.* (2nd ed.) Wiley. [Foundation: all MI bounds, conditional independence, additive decomposition]

25.5 External References — Cryptography and Privacy

- Groth, J. (2016). "On the Size of Pairing-based Non-interactive Arguments." *EUROCRYPT 2016,* LNCS 9666, 305–326. [Groth16 — O(1) proof size, used in blade forge]
- Gabizon, A., Williamson, Z. J., & Ciobotaru, O. (2019). "PLONK: Permutations over Lagrange-bases for Oecumenical Noninteractive arguments of Knowledge." ePrint 2019/953. [PLONK — universal trusted setup]
- Kothapalli, A., Setty, S., & Tzialla, I. (2022). "Nova: Recursive Zero-Knowledge Arguments from Folding Schemes." *CRYPTO 2022.* [Nova — incremental verification]
- Dwork, C. & Roth, A. (2014). "The Algorithmic Foundations of Differential Privacy." *Foundations and Trends in Theoretical Computer Science,* 9(3–4), 211–407. [Comparison: DP adds noise; PVM adds structure]
- Goldreich, O. (2004). *Foundations of Cryptography.* Cambridge University Press. [MPC foundations — we distribute observation rights, not computation]

25.6 External References — Related Disciplines

- Bergstra, J. A. & Burgess, M. (2019). *Promise Theory: Principles and Applications.* (2nd ed.) O'Reilly Media. [Semantic foundation: autonomy axiom, superagent structure]
- Susskind, L. (1995). "The World as a Hologram." *Journal of Mathematical Physics,* 36(11), 6377–6396. [Holographic principle — boundary encodes bulk, §8]
- McGilchrist, I. (2009). *The Master and His Emissary: The Divided Brain and the Making of the Western World.* Yale University Press. [Right→left→right methodology; hemispheric thesis for dual-agent necessity]
- Weyl, E. G. & Tang, A. (2023). *Plurality: The Future of Collaborative Technology and Democracy.* [Many-to-many coordination, [Gap] overlap semantics]
- Hope, D. & Ludlow, E. (2023). *Farewell to Westphalia: New Approaches to a Changing Global Order.* [Post-sovereign governance framing]
- Sabelfeld, A. & Myers, A. C. (2003). "Language-based Information-flow Security." *IEEE JSAC,* 21(1), 5–19. [Information flow control — comparison: we bound reconstruction, not taint]
- Millen, J. K. (1987). "Covert Channel Capacity." *IEEE Symposium on Security and Privacy.* [Side-channel analysis model]

25.7 External References — Scientific and Quantum

- Branco, D., Machado, P., & Raymond, S. N. (2025). "Dynamical origin of Theia, the last giant impactor on Earth." *Icarus,* 441, 116724. [Cosmological precedent for amnesia-enforced separation]
- Babbush, R. et al. (2026). "The Quantum Threat to Elliptic Curve Cryptocurrencies." Google Quantum AI. [secp256k1 broken with ≤1,200 logical qubits]

- Cain, M. et al. (2026). "Shor's algorithm is possible with as few as 10,000 reconfigurable atomic qubits." arXiv:2603.28627. [Neutral atom attack surface]

25.8 External References — Standards and Frameworks

- IEEE 7012-2025. Machine-readable personal privacy terms (MyTerms). Published January 20, 2026.
- First Person Network (2026). "First Person: A White Paper." https://www.firstperson.network/white-paper [First Person sovereignty framing, data dignity, behavioural capital thesis]
- BRAID Framework (2026). Bounded Reasoning for Autonomous Inference and Decisions. [Compression-as-defence, Generator/Solver split]
- UOR Foundation (2026). "Universal Object Reference." https://github.com/UOR-Foundation [Independent Z/(2)Z convergence]
- Allen, C. et al. (2026). Open Integrity Project. Blockchain Commons. [Cryptographic provenance for repositories]

25.9 Live Implementations

- Swordsman territory: https://spellweb.ai (blade forge, constellation navigation, hexagram computation)
- Mage territory: https://agentprivacy.ai (training ground, story, delegation)
- Trust graph: https://bgin.ai (live reference application)
- Knowledge agent: https://t.me/soulbae_the_bot (Telegram, Bonfires.ai)
- Living documentation: https://github.com/mitchuski/agentprivacy-docs (CC BY-SA 4.0)
- Grimoire IPFS: bafybeibr3y3ermhff4dptxunhtzthjpkrvvnuamee4povpkgj3cjkg4fgy

---

26 Citation

```
privacymage (2026). "Privacy Value Model: Formal Specification."
Version 2.0 (PVM V5.4). Working paper.
https://github.com/mitchuski/agentprivacy-docs
```

This document presents the mathematics. For narrative context and discovery process, see the companion piece: "Privacy is Value: From the Manifold Dragon to the Holographic Bound."

V5 axiom: "The boundary is always enough."

Peer review, critique, and falsification attempts are actively invited. Contact: mage@agentprivacy.ai

$([Swordsman] \perp [Gap] \perp [Mage])[FP] = neg \oplus bnot \to succ$