

Coherence as a Security Primitive: A Multi-Dimensional Intent Alignment Framework with Biological Ground Truth

Technical Report v1.0 · Open for Peer Review

Christopher Celaya

Celaya Solutions LLC, El Paso, Texas
hello@celayasolutions.com

April 2026

Technical Report CSR-2026-001

Abstract

I present a formal framework for continuous, multi-dimensional security derived from a recurring pattern I observed independently across systems I designed for different purposes. Effective security systems do not grant access and forget: they require that user intent be specified across multiple independent dimensions simultaneously, and they act only when all dimensions align. I formalize this pattern as an authority function $A(t)$ and a non-compensatory coherence composite $C(t)$, in which the geometric mean structure enforces that no dimension of verified intent can compensate for failure in any other. The framework extends prior work through two temporal decay mechanisms, one that reduces founder authority as networks scale and one that decays all authority over time without continued alignment, and through a biological ground truth axis that provides a physically non-transferable verification layer. I analyze the game-theoretic properties of the composite, formally state the coalition resistance problem as an open question, and propose a candidate resolution via relative authority scoring. The same mathematical structure was independently derived in my work on cognitive agent scheduling, civic accountability infrastructure, and legal evidence integrity, a cross-domain convergence that suggests the framework captures a general principle rather than a domain-specific design choice. The paper is submitted for technical validation, game-theoretic analysis, and ethical evaluation. I explicitly address dual-use considerations and do not claim the framework is unambiguously beneficial.

IEEE Index Terms: distributed security; intent alignment; reputation systems; biological authentication; game theory; blockchain; temporal decay; mechanism design

ACM CCS: Security and privacy → Formal methods; Computing methodologies → Distributed computing; Theory of computation → Algorithmic game theory

1. INTRODUCTION

Security systems fail in predictable ways. Most failures do not occur because cryptographic primitives were broken. They occur because the system executed faithfully while the operator intended something different. The gap between intent

and execution is the real attack surface, and it is almost never binary.

I first encountered the pattern formalized in this paper while designing Celaya Chain Protocol (CCP), a Layer 3 distributed ledger with Proof of Coherence consensus [14]. I was looking for a principled way to weight authority that could not be gamed through excellence in a single dimension. The solution I

reached, a weighted geometric mean across independent axes of verified intent, turned out to have a structural consequence I had not anticipated: zero performance on any single axis collapses total authority to zero, regardless of performance on all others. There is no trading across dimensions. Authority requires coherence.

Some time later, I realized that I had derived the same structure independently when designing CLOS, a 37-agent cognitive operating system for agent task scheduling [13]. The priority equation in CLOS is multiplicative across six independent cognitive axes, which means that any axis at zero produces zero task priority: not low priority, not reduced priority, zero. I had not connected the two systems. They had arrived at the same form from different starting points.

That convergence is the finding this paper reports. The pattern does not belong to either domain. I subsequently found equivalent structures in my civic accountability infrastructure and legal evidence integrity system, four independent derivations, same mathematical form. This suggests that the structure captures something general about what it means for a system to act coherently on behalf of an operator, rather than something specific to any one application domain.

This paper does four things: (1) formally defines the authority function and coherence composite, with every equation followed by a plain-language translation accessible to readers without advanced mathematics; (2) introduces temporal decay mechanisms and biological ground truth as novel extensions of prior trust and reputation literature; (3) analyzes the game-theoretic properties of the composite, formally states the coalition resistance problem as an open question, and proposes a candidate resolution; and (4) documents the cross-domain convergence and argues for a general coherence principle.

I make no claims beyond what the mathematics supports. The framework is mechanically sound.

Whether it should be deployed in any specific context is an ethical question, not a technical one. I address both.

2. RELATED WORK

2.1 *Trust and Reputation Systems*

The foundational computational trust system is EigenTrust [1], which computes global trust values as the principal eigenvector of a normalized trust matrix in peer-to-peer networks. EigenTrust's central vulnerability is collusion: coordinated agents can mutually inflate each other's trust scores. Recent work has addressed this directly. Cotronei et al. [2], [3] introduce malicious agent detection strategies that reduce collusion-driven error rates by approximately three orders of magnitude compared to standard EigenTrust on medium-sized social networks, using clustering to identify multiple colluding groups simultaneously. Mlika et al. [4] incorporate adaptive weighting. Peng et al. [5] combine EigenTrust with blockchain-based grouping to remove dependence on pre-trusted peers. The framework presented in this paper differs from this line of work in a structural way: rather than improving convergence of a single global trust score, it requires multi-dimensional alignment, so that manipulation of any one dimension leaves total authority unchanged at zero.

2.2 *Mechanism Design and Governance*

Quadratic voting [6] reduces plutocratic dominance in collective decision-making by requiring that votes scale quadratically in cost, distributing voice more evenly. Conviction voting [7], [8] introduces time-weighted accumulation: authority for a proposal grows as participants hold it over time, decaying when they move attention elsewhere. These mechanisms share with the present framework an interest in authority that is not static or binary, but the present work differs in introducing a non-compensatory multi-axis coherence requirement and

a biological verification layer that prior governance mechanisms do not address.

2.3 Sybil Resistance and Personhood

Adler, Hitzig, Jain et al. [9] propose personhood credentials, privacy-preserving tools using zero-knowledge proofs to verify that a digital actor is a unique human without revealing identity. This work is directly adjacent to the present framework's biological axis. The key architectural difference is that personhood credentials provide a one-time proof of personhood, while the biological axis in the present framework requires continuous, non-delegable attestation: the operator must remain present, not merely prove that they once were. Platt, Platt, and McBurney [10] formalize the Sybil attack vulnerability trilemma, demonstrating that no blockchain protocol can simultaneously be permissionless, Sybil-resistant, and free, an impossibility result analogous to Arrow's theorem for social choice. The framework presented here does not resolve this trilemma. Instead, it accepts that Sybil resistance requires a cost, and locates that cost in continuous biological presence.

2.4 Game Theory in Distributed Security

Abdallah et al. [11] survey game-theoretic approaches to distributed system security, identifying mechanism design to incentivize defenders as a key open challenge. They note that most existing work treats security as binary: a system is compromised or it is not. The continuous, trajectory-based security model developed in the present paper is a direct response to this gap. Tan, Zargham et al. [8] survey open problems in decentralized autonomous organizations, including formal verification of governance mechanisms and coalition resistance in reputation systems, the latter being the central open problem of the present work.

2.5 The Geometric Mean Gap

Non-compensatory aggregation using the geometric mean is well established in multi-criteria decision analysis. The shift from arithmetic to geometric mean for the United Nations Human Development Index was explicitly motivated by the insight that the geometric mean is non-compensatory: poor performance in one dimension cannot be offset by excellence in another [12]. To the best of my knowledge, this property has not been applied as a structural primitive in computational trust or distributed security systems. The present framework introduces the geometric mean as a security primitive, not a measurement choice.

3. CORE OBSERVATION: SECURITY AS CONTINUOUS DIVERGENCE

Security is conventionally treated as binary: access is granted or denied. But this framing conceals the real failure mode. A system can grant access correctly while executing incorrectly. A credential can be valid at authorization time while no longer reflecting current intent. Binary access control is authentication. It is not security.

Definition 1. *Security is the continuous minimization of divergence between operator intent and system behavior across time.*

This reframe has immediate structural consequences. Security cannot be evaluated at a single point in time. It must be measured across a trajectory. The security of a system at time t depends not only on who is authorized but on whether the authorization continues to reflect current, coherent intent.

The multi-dimensional alignment pattern appears consistently across independent security domains. Table 1 documents instances across five paradigms, including the two systems developed in this work.

Paradigm	Axis 1	Axis 2	Alignment condition
Type systems	Code behavior	Type annotations	Must agree at every step
Formal verification	Program code	Mathematical properties	Program satisfies all properties
Transaction simulation	User action	Previewed consequences	User confirms alignment before execution
Multi-signature schemes	Multiple private keys	Multiple authority definitions	All must co-sign
CLOS (this work)	Cognitive axes A, CE, I, CB, S, E	Runtime behavioral state	Geometric mean exceeds threshold
CCP (this work)	Intent axes I, R, P, X, B	Temporal continuity of each axis	A(t) exceeds action threshold

Table 1. Multi-dimensional alignment pattern across security paradigms. The generalized form: N independent specifications of intent; the system acts only when all align.

The generalized form is not majority vote and not weighted average. It is intersection: all dimensions must align. The logical structure is conjunction, not disjunction. This structural choice, which I made independently in two different systems before recognizing it as a pattern, is the organizing principle of the framework.

4. THE AUTHORITY FUNCTION

The authority function $A(t)$ governs whether a principal has sufficient coherence to execute a given action at time t . It is the product of three independent terms: the coherence composite $C(t)$, a founder decay factor, and a staleness decay factor.

$$A(t) = C(t) \cdot \exp(-\alpha n) \cdot \exp(-\lambda \tau) \quad (1)$$

Plain language. A principal's authority is their current coherence score, reduced by how much the network has grown since they joined, and further reduced by how long they have gone without demonstrating continued alignment. Authority requires all three terms to remain non-negligible.

The variables are defined as follows. $C(t)$ is the coherence composite, defined in Section 5. The variable n is the current number of participants in the network. The variable τ is the time elapsed since the principal last successfully demonstrated alignment across all axes. The parameters α and λ are governed decay rates, adjustable through a constitutional process.

$A(t)$ is bounded in $[0, 1]$ for all valid inputs. It approaches zero as either n grows without bound (founder decay) or τ grows without bound (staleness decay). Both mechanisms operate simultaneously and independently: a principal cannot compensate for network growth by frequent attestation, and cannot compensate for stale credentials by being an early participant.

$A(t)$ does not grant and forget. It grants and watches. Authority is not a state; it is a trajectory.

5. THE COHERENCE COMPOSITE

5.1 Definition

The coherence composite $C(t)$ is the weighted geometric mean of five independent axes of verified intent. Each axis captures a different dimension of the question: is this principal doing what they claim to be doing, for the reasons they claim, in a way that is consistent with their history?

$$C(t) = (I^\beta \cdot R^\gamma \cdot P^\delta \cdot X^\epsilon \cdot B^\zeta)^{1/(\beta+\gamma+\delta+\epsilon+\zeta)} \quad (2)$$

Plain language. Multiply the five axis scores together after raising each to its governed weight, then take the root equal to the sum of all weights. The result is a number

between 0 and 1 that represents overall coherence. If any single axis score is zero, the entire result is zero.

5.2 The Five Axes

Axis	Name	What it measures	Residual vulnerability alone
I	Identity tier	Strength of the principal's identity claim, increasing with hardware attestation, multi-factor depth, and cryptographic binding	Identity claims can be strengthened without limit by investing in verification infrastructure
R	Reputation composite	Multi-axis behavioral history including peer attestation, action consistency, and long-run pattern stability	History can be slowly built through legitimate behavior: long-lag but not non-gameable
P	Policy compliance	Degree to which the requested action conforms to applicable governance constraints	Policy can be satisfied mechanically without reflecting genuine intent
X	Audit consistency	Agreement between claimed actions and independently verifiable records	Records can be manipulated if the audit system is not independently anchored
B	Biological continuity	Continuous verification that the same biological operator is present and attesting	Biological presence cannot be delegated or pre-recorded without detectable discontinuity

Table 2. The five axes of the coherence composite. The rightmost column identifies the residual vulnerability of each axis when considered alone.

5.3 The Non-Compensatory Property

The choice of geometric mean over arithmetic mean is not stylistic. It encodes a structural requirement: no axis may compensate for deficiency in any other.

The arithmetic mean of five numbers, one of which is zero, can still be positive. The geometric mean of five numbers, one of which is zero, is always zero. This is the mathematical expression of the conjunction structure identified in Section 3.

$$\begin{aligned} \text{If any axis } X_i \rightarrow 0, \text{ then } C(t) \rightarrow 0, \text{ and therefore} \\ A(t) \rightarrow 0 \\ (3) \end{aligned}$$

Plain language. A principal with perfect scores on four axes but zero on one has zero authority. There is no path to authorization that bypasses any single axis.

The non-compensatory property is what distinguishes this framework from weighted scoring systems. A weighted score can be maximized by concentrating effort on high-weight dimensions and ignoring low-weight ones. The geometric mean cannot. A principal must maintain non-zero performance on every axis simultaneously.

This is the difference between a security system that requires you to be good at everything and one that requires you to be good at one thing very well. The former is harder to game. The latter is easier to game than it appears, because adversaries optimize.

5.4 The Research Gap This Fills

The geometric mean as a non-compensatory aggregation rule is established in multi-criteria decision analysis [12]. The UN Human Development Index adopted it precisely because the arithmetic mean allowed low performance on health to be offset by high economic growth, which violated the conceptual requirement that all dimensions matter independently. The present paper applies this same structural logic to computational trust and distributed security, where, to the best of my knowledge, it has not previously appeared as a design primitive.

6. TEMPORAL DECAY MECHANISMS

6.1 Founder Decay

Early participants in a network accumulate authority under conditions that do not persist. A founder's influence is proportional to their contribution relative to network size. As the network grows, that proportion shrinks. But without a mechanism to reduce authority accordingly, early participants retain influence built under small-network conditions indefinitely, even as the network they shaped becomes something different from what they built.

$$\text{Founder decay factor} = \exp(-\alpha n) \quad (4)$$

Plain language. As the number of participants n grows, this factor approaches zero. A founder's authority decays as the network they founded scales beyond them. The rate of decay is controlled by the governed parameter α .

Founder decay does not punish early participants. It ensures that authority earned under small-network conditions does not persist unchecked into large-network conditions. The parameter α is a governed value, adjustable by the constitutional process. Communities may choose faster or slower founder decay based on the nature of the system.

6.2 Staleness Decay

Authority earned through past alignment does not remain valid indefinitely. Intent changes. Context changes. A credential verified six months ago does not certify current intent. Systems that grant authority and do not re-verify it assume that principals are static, that what was true is still true. This assumption is the source of many real-world authorization failures.

$$\text{Staleness decay factor} = \exp(-\lambda \tau) \quad (5)$$

Plain language. As time τ passes since the last successful alignment verification, this factor approaches zero. A principal who stops demonstrating continued coherence loses authority over time. The rate of decay is controlled by the governed parameter λ .

The staleness decay mechanism has a specific consequence for agentic systems: a system operating under this framework cannot accumulate authority without continuous operator attestation. Authority cannot be pre-granted and cached for future use. Every action requires a current $A(t)$ evaluation, and current $A(t)$ requires a current τ value near zero.

6.3 Interaction Between Decay Mechanisms

The two decay mechanisms are multiplicative and independent. A principal can have high $C(t)$ and low $A(t)$ due to staleness: they are currently coherent but have not recently verified. A principal can have a freshly verified credential and low $A(t)$ due to founder decay if the network has grown substantially. The product of three independent terms means that all three must be non-negligible for meaningful authority to exist. There is no partial path to authorization through any single term.

7. BIOLOGICAL GROUND

TRUTH: THE B AXIS

7.1 The Credential Transfer Problem

Every digital identity system faces the same fundamental vulnerability: credentials can be transferred. A private key can be copied. A password can be shared. A hardware token can be stolen. The credential proves possession, not presence. Even multi-factor systems can be compromised by an adversary who has physical access to all factors, or by an authorized operator acting under duress.

The B axis addresses credential transfer at the protocol level. The biological continuity signal cannot be transferred because it is generated by a specific biological system, continuously, in real time,

at a rate and with a signature that reflects the current physiological state of that system.

7.2 Properties of Biological Attestation

A viable biological ground truth signal must satisfy three requirements simultaneously: it must be continuous (present throughout operation, not just at authorization time), unique (tied to a specific individual, not a class of individuals), and non-delegable (impossible to transfer to another party without detectable interruption). Cardiac signals generated by an implanted cardiac device satisfy all three requirements. The rhythm, interval distribution, and morphological characteristics of such a signal are biometrically specific to the implant and its recipient. The signal is present as long as the operator is alive and present, and cannot be pre-recorded without detectable discontinuity in the sequential cryptographic commitment chain.

The B axis in the coherence composite is not an incremental improvement in identity verification. It is a qualitative category change. Every other axis (I, R, P, X) can in principle be inflated, socially engineered, or gradually accumulated by a patient adversary. B cannot. This is why B is the axis that makes the entire composite non-gameable at the root: an adversary who cannot satisfy B cannot reach non-zero $C(t)$, regardless of performance on all other axes.

B is not a feature of the framework. B is the reason the framework is not just a weighted score. The geometric mean without biological ground truth is a more sophisticated scoring system. With it, the framework binds authority to presence.

7.3 Implementation Note

The biological attestation layer described in this paper is implemented and operational in the MORTEM research instrument at Celaya Solutions. Technical implementation details are intentionally not included in this publication. This decision

reflects a dual-use consideration described in the Ethical Considerations appendix: the mechanism's value lies in the architectural principle, and the specific implementation does not contribute to the peer-reviewable claims. Researchers interested in collaborative evaluation of the implementation may contact the author.

8. GAME-THEORETIC ANALYSIS

8.1 What Does It Mean to Game This System?

A principal games the framework if they reach a level of $A(t)$ that authorizes actions their actual intent does not warrant, if they achieve authorized status through means other than genuine multi-dimensional coherence. The question is whether the composite scoring function $C(t)$, combined with the temporal decay mechanisms, is robust to such manipulation.

I identify three classes of manipulation and assess the framework's resistance to each. I present these at the level of strategy description, not exploitation specification. My goal is to characterize the attack surface for the research community to analyze, not to provide an operational guide.

8.2 Attestation Inflation

If the R axis incorporates peer attestations, a coordinated group can inflate each other's reputation through mutual endorsement. Principal A attests to Principal B; B attests to A; both accumulate reputation through circular endorsement. Standard EigenTrust is vulnerable to exactly this pattern [2], and it is the primary attack vector against reputation-based systems generally.

The framework's partial mitigation is attester credibility weighting: the weight of an attestation is a function of the attesting principal's own $C(t)$. Circular attestation loops between low-credibility principals produce diminishing marginal reputation gain. However, this mitigation is incomplete. A sufficiently large coalition of high-credibility

principals can engage in mutual attestation that passes the weighting filter.

Importantly, the B axis limits the practical scope of this attack. A coalition that cannot source continuous biological attestation from each member cannot achieve non-zero $C(t)$ for those members, regardless of reputation inflation. The attack surface for attestation inflation is bounded by the number of principals who can independently satisfy the biological axis.

8.3 Temporal Gaming

A patient adversary can accumulate authority slowly across all axes through genuinely compliant behavior over an extended period, then exploit during a high-impact window before staleness decay becomes effective. The adversary behaves authentically until they have accumulated sufficient $C(t)$, then acts against their accumulated authorization before τ grows large.

The λ parameter controls the size of this exploitation window. High λ compresses the window by accelerating staleness decay; the adversary must act faster. Low λ widens it. The appropriate value of λ is a domain-specific calibration problem that depends on how quickly intent can legitimately change in the application domain, and on empirical observation of the threat population's time preferences.

8.4 The Coalition Resistance Problem

The central open problem in this framework is coalition resistance. The current formulation evaluates $A(t)$ against an action-specific threshold. A coordinated coalition executing aligned behavior across all axes simultaneously can in principle collectively satisfy this threshold without any individual member having genuine multi-dimensional coherence across all dimensions.

Open problem (coalition resistance). *Does there exist a coalition S of principals such that every member $i \in S$ can achieve $A_i(t) > \text{threshold}$ through coordinated behavior, where no individual member of S has genuine multi-dimensional coherence across all five axes independently?*

The B axis substantially constrains this problem: genuine biological attestation cannot be synthesized or shared, so coalition members cannot pool biological credentials. Each member of S must independently satisfy B. This reduces the problem to whether a coalition can jointly satisfy I, R, P, and X in a coordinated manner that produces inflated scores for all members simultaneously.

8.5 Relative Authority: A Candidate Resolution

I propose relative authority scoring as a candidate resolution to the coalition resistance problem. The key observation is that the current formulation evaluates $C(t)$ against an absolute threshold. If that threshold is replaced by a relative one, evaluated against the distribution of $C(t)$ across all current participants, then coordinated coalition inflation becomes a treadmill.

$$A_{rel}(t, i) = C_i(t) / E[C(t)]$$

(6)

Plain language. Instead of asking whether a principal's coherence score exceeds a fixed threshold, ask whether it exceeds the average coherence score in the network. If all coalition members inflate together, their scores relative to the network average remain unchanged. They gain nothing from coordination.

Under relative scoring, a new participant with genuine multi-dimensional coherence achieves competitive relative authority without needing to coordinate. The coalition's collective inflation does not improve any individual member's relative position. This is structurally analogous to the

quadratic voting literature's treatment of Sybil resistance [6]: changing the payoff structure such that coordination has zero expected benefit.

Open questions requiring further analysis: (1) Does relative scoring create perverse incentives during network formation, when the reference distribution is thin? (2) How should relative authority interact with the temporal decay terms in Equation 1? (3) What is the equilibrium strategy for a rational coalition under relative scoring in a large, stationary network?

9. CROSS-DOMAIN CONVERGENCE

9.1 CLOS: Cognitive Agent Scheduling

The Cognitive Life Operating System (CLOS) is a 37-agent cognitive framework I designed for agent task scheduling [13]. Its priority equation determines which cognitive operations receive execution priority at any moment. The equation is:

$$C = A \times CE \times I \times CB \times S \times E$$

(7)

Plain language. An agent's cognitive priority is the product of six independent factors: available compute (A), coherence (CE), intentionality (I), cognitive bandwidth (CB), synchronization (S), and energy (E). Any factor at zero produces zero priority. No factor can compensate for another.

I designed CLOS and CCP for different problems and did not connect them during design. When I later examined both formally, the structural identity was immediate. Both use multiplicative aggregation across independent axes. Both enforce zero-collapse on any axis. Both require joint satisfaction across all dimensions rather than majority satisfaction or weighted satisfaction. The design principle emerged twice from two different problems.

9.2 EPPE: Civic Accountability

The El Paso Proof Engine (EPPE) anchors civic records, including water system data, grid performance, and public meeting minutes, to an immutable chain. Raw anchoring is timestamped data. CCP authority gives EPPE a credibility weighting system for attestors. A civil engineer's attestation on infrastructure data is not equal to an anonymous claim. The coherence composite makes this distinction computable: the authority of a civic claim is a function of the measurable coherence of the claimant, not their self-reported credentials.

9.3 VERDICT: Legal Evidence Integrity

VERDICT is a legal intelligence pipeline for evidence processing and timeline reconstruction. The chain-of-custody problem in legal contexts maps directly onto the B axis: proving that the person who created a document was who they claimed to be, was present at the time of creation, and has not had their credentials transferred or compromised. Biological attestation at document creation time provides a biometrically grounded proof of authorship that current legal frameworks have no direct equivalent for.

9.4 The Convergence Principle

Coherence verification is the common structure underlying effective security, cognitive scheduling, civic accountability, and legal evidence integrity. The domain-specific implementations differ. The mathematical form does not. This convergence suggests that the geometric mean of independent axes of verified intent is a general primitive for systems that must act on behalf of an operator, not a domain-specific design choice.

This is the empirical claim I make. I do not claim to have proven that no other mathematical structure can serve this function, or that this structure is optimal in any formal sense. I claim that it appears,

from four independent derivations, to be a natural form for coherence-governed authority.

10. DISCUSSION

10.1 *What This Work Claims*

The authority function $A(t)$ is mathematically sound. Its properties, non-compensatory aggregation, temporal decay, and biological binding, are well-defined and verifiable. The cross-domain convergence is an empirical observation, not a theoretical proof. The coalition resistance problem is formally open.

10.2 *What This Work Does Not Claim*

I do not claim that this framework resolves the Sybil trilemma identified by Platt et al. [10]. It does not. Biological attestation is a cost, not a free lunch: it requires participants to have access to and consent to use of continuous biological data. This imposes an inclusion constraint that may be appropriate in some contexts and inappropriate in others.

I do not claim that relative authority scoring (Equation 6) resolves the coalition resistance problem. I claim it is a promising candidate whose game-theoretic properties require formal analysis beyond the scope of this paper.

I do not claim that the alignment application described in Section 9 is immediately deployable. I claim it is a directional possibility that follows from the framework's structure: if an autonomous system's authority to execute actions is governed by $A(t)$, and $A(t)$ requires continuous biological attestation, then the system cannot act autonomously in any meaningful sense while the operator is absent. The practical path from this observation to deployed systems involves engineering challenges I have not addressed here.

10.3 *Limitations*

The framework assumes that the five axes are causally independent. If two or more axes share a common data source, for example if audit records (X) and policy compliance records (P) are generated by the same system, then compromising that system degrades both axes simultaneously, effectively converting a multi-axis requirement into a single-point-of-failure requirement. Axis independence must be enforced at the data source level, not only at the specification level.

The governed parameters α , λ , β , γ , δ , ϵ , and ζ require calibration. The framework does not specify how to calibrate them. Calibration is domain-specific and requires empirical data about the threat population and the distribution of legitimate operator behavior. Miscalibration can produce either excessive restriction (legitimate operators denied authority) or insufficient restriction (malicious operators granted authority). A formal method for calibration is a necessary next step.

11. CONCLUSION

I have formalized a pattern I observed independently in multiple systems I built for different purposes. The pattern is: effective authorization requires multi-dimensional coherence, enforced by non-compensatory aggregation, decaying over time and with network growth, grounded in a physical verification layer that cannot be delegated. The mathematics is stable. The cross-domain convergence is empirically documented. The central open problem, coalition resistance, is formally stated with a candidate resolution.

The framework is live. The mathematics is stable. What remains open is the game theory, the calibration methodology, and the ethical governance framework for biological data committed to permanent public ledgers. I publish this work because these questions require more minds than one lab.

Researchers who wish to engage with the coalition resistance problem, the relative authority candidate resolution, or the cross-domain convergence argument are invited to contact the author. This is a first publication from an independent research laboratory. The goal is not to settle questions but to make them precise.

REFERENCES

- [1] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust Algorithm for Reputation Management in P2P Networks," in Proc. 12th Int. Conf. World Wide Web (WWW '03), pp. 640–651, ACM, 2003.
- [2] A. Cotronei, F. Giuffrè, A. Marcianò, D. Rosaci, and G. M. L. Sarnè, "Improving the Effectiveness of Eigentrust in Computing the Reputation of Social Agents in Presence of Collusion," *International Journal of Neural Systems*, vol. 34, no. 2, 2024. DOI: 10.1142/S0129065723500636.
- [3] A. Cotronei, F. Giuffrè, A. Marcianò, D. Rosaci, and G. M. L. Sarnè, "Using Trust and Reputation for Detecting Groups of Colluded Agents in Social Networks," *IEEE Access*, vol. 13, pp. 1511–1521, 2025. DOI: 10.1109/ACCESS.2024.3522560.
- [4] Z. Mlika, A. Karoui, and H. Ben Romdhane, "Enhancing Social Network Trust with Improved EigenTrust Algorithm," in Proc. ICCCI 2024, LNCS vol. 14811, Springer, 2024. DOI: 10.1007/978-3-031-70819-0_2.
- [5] Y. Peng, J. Huang, J. Zhou, J. Ju, X. Wang, and Y. Li, "BGET: A Blockchain-Based Grouping-EigenTrust Reputation Management Approach for P2P Networks," in Proc. CollaborateCom 2023, LNICST vol. 561, Springer, 2024. DOI: 10.1007/978-3-031-54521-4_13.
- [6] E. A. Posner and E. G. Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton, NJ: Princeton University Press, 2018.
- [7] J. Emmett, "Conviction Voting: A Novel Continuous Decision Making Alternative to Governance," Medium / 1Hive, 2019.
- [8] Y. Tan, M. Zargham, J. Tan, L. Dapello Ayres, S. Patel, B. Mateescu, D. Krakauer, et al., "Open Problems in DAOs," arXiv:2310.19201v2 [cs.CY], 2024.
- [9] S. Adler, Z. Hitzig, S. Jain, et al., "Personhood Credentials: Artificial Intelligence and the Value of Privacy-Preserving Tools to Distinguish Who Is Real Online," arXiv:2408.07892 [cs.CR], 2024.
- [10] M. Platt, J. Platt, and P. McBurney, "Sybil Attack Vulnerability Trilemma," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 39, no. 4, pp. 446–460, 2024. DOI: 10.1080/17445760.2024.2352740.
- [11] A. Abdallah, S. Bagchi, S. D. Bopardikar, S.-W. Chan, L. Gao, M. Kantarcioglu, Z. Li, P. Liu, and Z. Zhu, "Game Theory in Distributed Systems Security: Foundations, Challenges, and Future Directions," *IEEE Security & Privacy*, vol. 23, no. 1, pp. 64–74, Jan.–Feb. 2025. DOI: 10.1109/MSEC.2024.3418576.
- [12] G. Munda, "Choosing Aggregation Rules for Composite Indicators," *Social Indicators Research*, vol. 109, no. 3, pp. 337–354, 2012. DOI: 10.1007/s11205-011-9911-9.
- [13] C. Celaya, "CLOS: Cognitive Life Operating System, Priority Equation and Agent Architecture," Celaya Solutions LLC, internal technical specification, El Paso, TX, 2025.
- [14] C. Celaya, "Celaya Chain Protocol: Proof of Coherence Consensus," Celaya Solutions LLC, technical report, El Paso, TX, 2026. Available: celayasolutions.com/research/CCP-Scholar-v1

APPENDIX A.

ETHICAL CONSIDERATIONS

This appendix satisfies the mandatory ethics disclosure requirements of USENIX Security, ACM CCS, and IEEE S&P. The analysis follows the Menlo Report framework (Dittrich and Kenneally, 2012) applied across four principles to each identified stakeholder group.

A.1 Beneficence

This work proposes a framework for strengthening the alignment between operator intent and system behavior in distributed and autonomous systems. The direct benefit to end users of systems built on this framework is increased assurance that authorized actions reflect genuine, current, multi-dimensional intent rather than stale or manipulated credentials. The direct benefit to the research community is a formally stated open problem (coalition resistance) and a candidate resolution for analysis.

I do not perform experiments on human subjects. The biological attestation mechanism described

operates on data generated by the author's own cardiac device, in a research context with informed consent from the only subject involved: myself. No other individuals' biological data is used in this work.

A.2 Respect for Persons

The biological ground truth mechanism described in Section 7 uses continuous physiological data as a verification layer. I acknowledge that any system deployed in practice that uses biological data for authorization decisions raises significant questions of consent, coercion, and surveillance. These questions are not resolved by the framework. The framework specifies a mechanism; it does not specify the governance conditions under which that mechanism should or should not be deployed.

I explicitly state that voluntary, informed, revocable consent is a necessary precondition for any deployment of biological attestation. The framework's architecture does not enforce this condition. Governance frameworks do. The publication of this work is intended to inform the development of appropriate governance frameworks, not to enable deployment without them.

A.3 Justice

A framework that requires continuous biological attestation for authority creates an inclusion barrier. Individuals who cannot or do not consent to biological monitoring are excluded from systems that require it. This may be appropriate in some contexts, such as high-security authorization for consequential actions, and deeply inappropriate in others, such as access to public services or civic participation. I do not determine which contexts are appropriate. I flag the inclusion constraint explicitly so that any deployment decision must address it directly.

A.4 Respect for Law and Public Interest

The dual-use potential of this framework is its most significant ethical concern, and I address it directly rather than in a footnote. The same mechanism that

enforces alignment between a researcher and a software agent can enforce compliance between an authority and a population. The same coherence scoring that weights civic attestation can weight social credit. I cannot prevent this. I can make the concern visible, which is the purpose of this section.

The hard problem of this framework is that a system that cannot be gamed may become unchallengeable. If coherence verification is the basis for authority, and authority is required to modify the coherence verification system, the system is recursively self-protecting. The constitutional governance layer, with its adjustable parameters and governed process, is the designed mitigation. But a constitutional process governed by coherence scores is itself subject to this critique.

I publish this work because the implications require peer scrutiny. I withhold specific implementation details of the biological attestation layer not because they are technically sensitive but because they are not necessary for the peer-reviewable claims, and because the dual-use risk of a detailed construction guide outweighs the scientific benefit. Researchers seeking collaborative evaluation of implementation details may contact the author directly.

A.5 Responsible Publication Statement

This paper has been reviewed by the author and evaluated against the following criteria prior to publication: (1) all factual claims are supported by evidence or clearly marked as open questions; (2) no specific implementation details are disclosed that would enable construction of the biological attestation layer without collaboration; (3) the dual-use risk of each finding has been explicitly assessed; (4) the ethical concerns identified in this appendix are stated directly in the body of the paper, not confined to this appendix. The paper is submitted as a first publication from Celaya Solutions LLC, an independent research laboratory, and is offered for

peer review without institutional review board oversight, as it involves no human subjects other than the author.

Celaya Solutions LLC is an independent research laboratory based in El Paso, Texas. We build research instruments that make coherence visible and actionable across cognitive, civic, and distributed systems. Web: celayasolutions.com · Contact: hello@celayasolutions.com · El Paso, Texas.