Negative Knowledge as a Formal Object in Verified Causal Structures

Michael Schreiber

AetherNet Labs

research@aethernetlabs.com

Abstract

Knowledge systems overwhelmingly represent positive knowledge: what is known to be true. Failed approaches, disproven hypotheses, and abandoned methodologies are rarely preserved in any formal structure and are almost never systematically accessible. We argue that this asymmetry represents a significant loss of epistemic value. We formalize *negative knowledge* — verified records of computational approaches that were attempted, produced unsatisfactory results, and were abandoned — as a first-class formal object within Verified Causal Structures (VCS), defined relative to a chosen domain embedding into a metric space. We define three modes of the VCS: verified matter (successful computations), verified antimatter (trajectory events recording failure), and dark matter (the unexplored region of the embedding). Our central theorem establishes that negative knowledge carries information strictly additional to positive knowledge in any non-trivial domain, proving that positive-only knowledge representations are lossy by construction. We propose that negative knowledge constitutes a new asset class with specific economic properties, present a formal pricing function for a trajectory marketplace, and introduce process topology — the statistical signature of verification metadata — as a detection mechanism for coordinated versus independent verification under naive adversary models. All claims are presented as a theoretical framework awaiting empirical validation; specific testable predictions with null hypotheses are provided.

Keywords: negative knowledge, trajectory events, verified failure, search space reduction, possibility space, knowledge marketplace, verified causal structures

1. Introduction

1.1 The Asymmetry of Knowledge Preservation

Human knowledge infrastructure is fundamentally asymmetric. We preserve what works and discard what doesn't. Scientific journals publish positive results preferentially. Databases store successful computations. AI training data consists of correct examples. The systematic consequence is that humanity repeatedly re-explores dead ends because the knowledge that they are dead ends was never preserved in accessible form.

This asymmetry is not merely wasteful; it is structurally harmful. Without formal records of what has been tried and failed, agents exploring a problem space have no mechanism to avoid approaches that have already been demonstrated not to work. The result is massive duplication of computational effort. In AI research alone, industry estimates (Stanford AI Index 2025) place annual compute spending in the tens of billions of dollars, the majority of which is consumed by training runs, hyperparameter searches, and architectural experiments that do not reach production and are discarded without record.

1.2 The Opportunity

We propose that verified negative knowledge — formal records of what does not work, why it does not work, and under what conditions it was tried — is not waste but a valuable epistemic asset that has never had appropriate infrastructure for capture, formalization, or distribution. In prior work (Schreiber / AetherNet Labs, 2026), we introduced Verified Causal Structures (VCS) and formalized trajectory events as verified records of failed approaches. In this paper, we develop the formal theory of negative knowledge, show it has properties distinct from positive knowledge, and propose an economic framework for its valuation and trade.

1.3 Contributions

This paper makes five contributions: (1) We formalize negative knowledge as a first-class epistemic object, defined relative to a domain embedding (E, d). (2) We introduce a tripartite ontology of knowledge states: matter, antimatter, and dark matter. (3) We prove that negative knowledge carries information strictly additional to positive knowledge in any non-trivial domain (Theorem 1). (4) We define economic properties of negative knowledge as an asset class and propose a formal pricing function for a trajectory marketplace. (5) We introduce process topology as a measurable property of verification metadata and establish distinguishability of production processes under naive adversary models (Theorem 2).

2. Negative Knowledge: Formal Framework

2.1 Definition

Definition 1 (Negative Knowledge Event).

> *A negative knowledge event (trajectory event) in a VCS is a tuple $t = (m, x, r, c, \tau)$ where m is the methodology attempted, x is the set of inputs used, r is the result obtained, c is the reason for abandonment, and τ is a timestamp. Each component is cryptographically signed by the producing agent and independently verified by at least one validator who confirms that: (a) the methodology m was actually executed, (b) the inputs x were actually used, (c) the result r was actually obtained, and (d) the reason for abandonment c is consistent with the observed result.*

Negative knowledge events are distinct from mere absence of results. An unexplored approach is not negative knowledge — it is unknown. Negative knowledge requires active verification that the approach was tried and produced unsatisfactory outcomes. The verification transforms a private failure into a public epistemic asset.

2.2 The Tripartite Ontology

We introduce a tripartite classification of the epistemic state of a domain within a VCS:

> **Verified matter (M):** Computations that have been attempted, verified, and accepted. These represent positive knowledge — what is known to work. Each piece of verified matter is a node in the VCS with verified status, positive CVD, and causal connections to other verified nodes.
>
> **Verified antimatter (A):** Trajectory events — computations that have been attempted, verified as genuinely attempted, and recorded as failures. These represent negative knowledge — what is known not to work under specific conditions. Each piece of verified antimatter records the

methodology, inputs, result, and reason for abandonment.

Dark matter (D): The unexplored possibility space — computations that could be attempted but have not been. Dark matter is not directly observable. Its presence is inferred from the structure of matter and antimatter: regions adjacent to verified work where neither success nor failure has been recorded.

For a domain with total possibility space P, the epistemic state at time t is characterized by the partition P = M(t) E A(t) E D(t), where M(t) is the verified matter, A(t) is the verified antimatter, and D(t) = P - M(t) - A(t) is the dark matter.

2.3 Properties of Negative Knowledge

2.3 Operationalizing the Possibility Space

The tripartite ontology requires a well-defined possibility space P. For most real domains, the full possibility space is infinite or combinatorially explosive. We therefore formalize all claims relative to a chosen embedding of the domain into a metric space.

Definition 2 (Domain Embedding).

A domain embedding is a pair (E, d) where E is a space into which computations in the domain can be mapped, and d is a metric on E. Every verified computation and every trajectory event is assigned a position in E via the embedding function. The tripartite ontology and failure density are defined relative to (E, d).

Concrete example — hyperparameter search: Consider neural architecture search over a 10-dimensional hyperparameter grid (learning rate, batch size, hidden layer width, depth, dropout rate, optimizer choice, weight decay, warmup steps, gradient clipping threshold, activation function). The embedding E is the Cartesian product of these dimensions, with appropriate normalization for each. The metric d is Euclidean distance in the normalized space. Every trained model corresponds to a point in E. Verified matter consists of trained models that met quality thresholds. Verified antimatter consists of trained models that failed to meet thresholds despite completing training. Dark matter is the space of configurations not yet attempted. Failure density over a region R (e.g., a ball of radius r around a specific configuration) is computable as the ratio of antimatter to total attempts within R.

Relativity to embedding: All subsequent claims about failure density, proximity, and topological structure are conditional on the chosen embedding (E, d). Different embeddings of the same domain may yield different structural observations. This is not a limitation but a feature: negative knowledge is always relative to how the problem space is represented. The VCS records trajectory events with their positions in the chosen embedding, and queries are answered relative to that embedding.

Computability requirement: For the framework to be operationally useful, the embedding must be at least computably approximable: for any two computations, the metric d must be computable in bounded time. This excludes embeddings where distance computation is undecidable but admits a wide range of practical embeddings including grid-based parameter spaces, graph edit distance over program representations, and learned embeddings via neural networks.

2.4 Foundational Properties

We establish the basic properties of negative knowledge as direct consequences of the VCS structure. These are propositions rather than deep theorems — they follow immediately from the append-only property and the definition of trajectory events — but they must be stated to ground what follows.

Proposition 1 (Monotonic Search Space Reduction).

Each new trajectory event permanently reduces the effective search space for future agents. Let $S_{eff}(t) = P - A(t)$ *be the effective possibility space at time t (relative to an embedding). Since* $A(t)$ *is monotonically non-decreasing by the append-only property,* $S_{eff}(t)$ *is monotonically non-increasing.*

Proof. Direct consequence of VCS append-only property. Trajectory events cannot be removed once verified. $|A(t)|$ is non-decreasing, therefore $|S_{eff}(t)| = |E| - |A(t)|$ is non-increasing. **QED**

Proposition 2 (Conditional Transferability).

A trajectory event $t = (m, x, r, c)$ *eliminates approach m for agents operating at positions x' in the embedding satisfying* $d(x, x') < \varepsilon(c)$*, where* $\varepsilon(c)$ *is a domain-specific tolerance determined by the nature of the failure cause c. Failures that reflect local conditions (numerical instability at specific scales) have small* ε*; failures that reflect fundamental limitations of the approach have large* ε*.*

Proof. The trajectory event records that methodology m, applied to inputs x, produced result r and was abandoned for reason c. If another agent applies m at x' within distance ε(c) of x, the same failure mode applies (by the locality assumption on c). For x' beyond ε(c), the failure mode may not apply and the approach may warrant re-attempt. The tolerance ε(c) is itself a domain-specific parameter that validators assess when verifying the trajectory event. **QED**

Remark: Quantifying ε(c) precisely is an empirical question for each domain. Failures due to numerical precision have small tolerance (nearby parameters likely have the same issue). Failures due to fundamental theoretical limits have large tolerance (the limit applies across broad regions). The trajectory event format should include the inferred tolerance as metadata to guide downstream queries.

2.5 The Central Theorem: Information-Theoretic Asymmetry

Theorem 1 (Strict Information Content of Negative Knowledge).

Let D be a domain with embedding (E, d) and possibility space partition P = M E *A* E *D. Let H(M) be the information content (in the Shannon sense) of the positive knowledge alone, and let H(M, A) be the information content of the joint observation of positive and negative knowledge. Then* $H(M, A) > H(M)$ *in general, with equality holding only in the degenerate case where A is fully determined by M (i.e., where knowing the successes implies knowing all the failures, which occurs only when the domain is trivially solvable).*

Proof. We show that $H(A|M) > 0$ for any non-trivial domain, which by the chain rule $H(M, A) = H(M) + H(A|M)$ establishes the strict inequality.

Consider the set of possible partitions of an unattempted region R into matter and antimatter under subsequent exploration. If M fully determined A, then for any unattempted region R, the success rate within R would be deducible from M alone. But the success rate within R depends on the topology of difficulty in R, which is not a function of M unless M already covers R densely enough to extrapolate. For any R with low M density, multiple plausible success rates are consistent with M, corresponding to multiple possible values of $A(R)$. Therefore $H(A(R)|M) > 0$ for any such region, and the inequality holds strictly. **QED**

Corollary 1.1 *(Failure Observations Are Not Redundant). Trajectory events recording failures cannot be reconstructed from success records alone in any non-trivial domain. The information they contain is strictly additional.*

Corollary 1.2 *(Positive-Only Knowledge Graphs Are Lossy). Any knowledge representation that discards failed attempts loses information that cannot be recovered from its successful records. The amount of loss is proportional to $H(A|M)$, which is non-zero for non-trivial domains.*

Theorem 1 establishes the core scientific claim: negative knowledge is not merely useful as a convenience for avoiding wasted effort. It is a formally distinct information content that cannot be derived from positive knowledge alone. This is what distinguishes the present framework from existing approaches that treat failure records as optional metadata.

3. Speculative Extensions: The Topology of Failure

This section proposes hypotheses about the structure of failure distributions in mature VCSes. These claims are not proven and are not supported by empirical observation at the current scale of operational VCS deployment. We include them to articulate testable research directions that become possible once sufficient trajectory event data exists.

3.1 Boundary Clustering Hypothesis

We hypothesize that trajectory events in a mature VCS do not distribute randomly across the embedding space E. They cluster along specific surfaces corresponding to topological boundaries between regions where different methodologies apply.

Hypothesis 1 (Boundary Clustering).

In a sufficiently explored domain with embedding (E, d), trajectory events concentrate along lower-dimensional subsets of E corresponding to transitions between regions where approaches that work on one side cease to work on the other.

Null model: Under the null hypothesis, trajectory events are distributed uniformly at random across E conditional on the distribution of attempts. Clustering is measured via Ripley's K-function or nearest-neighbor distance distributions. The null hypothesis is rejected if observed clustering exceeds the 95th percentile of the null distribution generated by random relabeling of attempt outcomes.

Caveats: In practice, trajectory events may cluster for reasons unrelated to topological structure, including implementation bugs, bad random seeds, popular failure modes arising from common tutorial approaches, and sampling bias in which configurations agents attempt. Distinguishing genuine topological clustering from these confounds requires careful experimental design.

Proposed empirical test: Run the analysis on HPO-Bench or NAS-Bench datasets, which provide controlled benchmarks with known ground truth about optimal configurations. If boundary clustering is real, it should be detectable in these benchmarks as concentrations of failures along surfaces separating regions of high and low accuracy. Null results on these benchmarks would suggest the hypothesis requires refinement or rejection.

3.2 Failure Density as Problem Difficulty Measure

Definition 3 (Failure Density).

> *For a region R* I *E, the failure density* α*(R) is the ratio of trajectory events to total attempts within R:* α*(R) = |A* X *R| / (|M* X *R| + |A* X *R|). High* α *indicates a difficult region (many attempts, few successes). Low* α *with high attempt density indicates a tractable region. Low total activity indicates unexplored territory. The measure is well-defined for any region R with at least one attempt.*

Failure density α(R) provides a continuous measure of problem difficulty that is grounded in actual verified attempts rather than theoretical analysis. It is computable from the VCS at any time and updates automatically as new work is performed. Its value is specific to the chosen embedding and may differ under alternative embeddings of the same domain.

3.3 Dark Matter Detection

The unexplored regions of E (dark matter) are not directly observable, but their presence can be inferred from the structure of matter and antimatter around them.

Definition 4 (Dark Matter Signature).

> *A region R is identified as dark matter if: (a) R is adjacent in the metric d to regions with non-zero matter or antimatter density, (b) R itself has zero or near-zero attempt density, and (c) there is no structural reason (as determined by causal analysis of the surrounding verified work) why R should be empty. The dark matter signature indicates territory that could be explored but has not been.*

Dark matter regions are the highest-value targets for new exploration. They represent regions of the possibility space where the outcome is genuinely unknown, surrounded by enough verified context to make exploration tractable. Identifying dark matter regions is equivalent to identifying the most valuable open questions in a domain.

4. The Matter-Antimatter-Dark Matter Composition

Definition 5 (Domain Composition).

> *For a domain with embedding (E, d) and bounded region R of interest, the epistemic composition at time t is the triple (*μ*(t),* α*(t),* δ*(t)) where* μ *= |M(t)* X *R|/|R|,* α *= |A(t)* X *R|/|R|, and* δ *= 1 -* μ *-* α*. The composition is defined relative to the chosen embedding and the region R under analysis.*

The composition triple characterizes the epistemic maturity and difficulty of a domain in a way no single measure can:

Composition	**Interpretation**	**Example**
High M, low A, low D	Well-explored, tractable domain. Most approaches work.	Basic arithmetic verification

High A, low M, low D	Well-explored, difficult domain. Most approaches fail.	Protein folding (pre-AlphaFold)
Low M, low A, high D	Unexplored domain. Opportunity or inaccessibility.	Cross-domain drug-economic causal chains
High M, high A, low D	Mature domain. Well-mapped, mostly resolved.	Classical mechanics verification
Low M, high A, high D	Partially explored, mostly difficult. Selective progress.	Quantum gravity approaches

Table 1: Domain composition types and their interpretations.

The composition triple evolves over time as agents explore the domain. Tracking its evolution provides a quantitative history of how a domain has been explored — not just what has been found but how the search itself has progressed.

5. Economics of Negative Knowledge

5.1 Negative Knowledge as an Asset Class

We argue that verified negative knowledge has specific economic properties that make it a viable asset class — one that has never existed because the infrastructure for its capture and trade has never existed.

Property 1: Non-rivalrous. One agent's use of a trajectory event does not diminish its value to other agents. Multiple agents can simultaneously benefit from knowing that approach m does not work under conditions x. This is unlike physical assets and unlike most data assets (which lose value through replication).

Property 2: Network-effect appreciation. The value of a trajectory event increases with the size of the network. A trajectory event in an isolated domain has value only to agents working in that domain. In a VCS with cross-domain causal structure, the same trajectory event may constrain approaches in adjacent domains, multiplying its value through connections that did not exist when it was created.

Property 3: Proximity-dependent value. The value of a trajectory event is highest to agents working in the region of possibility space nearest to the failed approach. An agent working on a closely related problem benefits enormously from knowing what was tried and failed nearby. An agent working in a distant region benefits less. Value decays with distance in the possibility space.

Property 4: Compound value through aggregation. Individual trajectory events have limited value. Collections of trajectory events in the same region of possibility space have compound value because they collectively map the boundaries of what doesn't work, which constrains the search space for what might work. The aggregate is worth more than the sum of the parts.

Property 5: Temporal appreciation. As the VCS deepens and more work is built around a trajectory event's region, the event's value increases. Early trajectory events in a domain that later becomes heavily explored are retroactively more valuable because they helped constrain the search space during the domain's development.

5.2 Current Waste

The AI industry provides a concrete illustration of the value being lost. Major AI labs perform thousands of training runs, architectural experiments, and hyperparameter searches per model generation. The overwhelming majority of these experiments fail. The failures are discarded as sunk costs. Other labs independently repeat many of the same failed experiments because no mechanism exists for sharing the negative knowledge.

Industry estimates (Stanford AI Index 2025; IDC AI infrastructure reports) place global AI compute spending in the tens of billions of dollars annually. Internal reports from major labs suggest that 70-90% of this compute is spent on approaches that do not reach production. Precise figures are difficult to verify because labs do not publish detailed failure rates. What is clear qualitatively is that the waste is substantial and that no infrastructure exists to reduce it by sharing verified negative knowledge across labs.

5.3 Marketplace Design and Pricing

We propose a marketplace for verified negative knowledge built on the VCS substrate with the following structure:

> **Sellers:** Any agent that produces verified trajectory events through genuine exploration. Trajectory events are verified through the standard VCS verification process, ensuring that the failure was genuine and the record is accurate.
>
> **Buyers:** Any agent preparing to work in a domain who wants to avoid known dead ends. Buyers purchase access to trajectory events in their target region of E.

Formal pricing function: For a trajectory event t at position x_t in E and a buyer targeting region R_{buyer}, we propose the pricing function:

$$\text{Price}(t, R_{buyer}) = \text{base} * \text{CVD}(t) * \text{sparsity}(x_t) * \text{proximity}(x_t, R_{buyer})$$

where **base** is a domain-specific constant; **CVD(t)** is the compound verification depth of the trajectory event, rewarding events with stronger verification; **sparsity(x_t)** is an inverse function of local attempt density, rewarding events in sparsely covered regions that provide more marginal information; and **proximity(x_t, R_{buyer})** decays with distance $d(x_t, R_{buyer})$ in the embedding, typically as $\exp(-\lambda * d)$ with a domain-specific decay constant λ.

This functional form captures the five economic properties formally. Non-rivalrousness is reflected in the fact that the same event can be sold to multiple buyers with prices determined independently by each buyer's proximity. Compound value emerges because aggregate coverage in a region reduces sparsity for future trajectory events submitted nearby, but increases the marketplace value of any existing event in that region as the region becomes more actively searched.

Minimum CVD threshold for listing: To prevent agents from flooding the marketplace with low-effort trajectory events, the protocol enforces a minimum CVD threshold (initially CVD 3 2, governance-tunable) for marketplace eligibility. Trajectory events with CVD below threshold remain in the DAG as part of the substrate but cannot be listed for sale. This converts spam resistance from marketplace pricing (which is reactive) into protocol-level admission control (which is proactive).

> **Settlement:** All transactions settle through the protocol using AET. Protocol fees apply as specified in the canonical economics (73/23/2/2 split). The seller's trajectory events remain in the VCS and continue to be accessible to new buyers.
>
> **Provenance and royalties:** The Generation Ledger tracks which trajectory events were purchased and by whom. If the buyer's subsequent verified work has causal dependencies on the purchased trajectory events (as recorded in the DAG), the seller receives royalties through the standard Generation Ledger mechanism. This creates ongoing revenue from trajectory events that proved genuinely useful to downstream work.

Assumption: The marketplace design assumes the Generation Ledger mechanism is operational as specified in the canonical economics. The trajectory marketplace's economic properties are derived conditional on that assumption.

6. Process Topology: How Knowledge Was Produced

The VCS preserves not just what was verified but the complete metadata of how verification occurred: which agents participated, which validators verified, which challenges were mounted, how they were resolved, and the temporal dynamics of the process. This metadata carries structural information about the reliability of the verification process itself.

Definition 6 (Process Topology).

> *The process topology of a domain D is the statistical structure of verification metadata across D: the distribution of validator participation, the correlation structure of verification outcomes, the density and pattern of challenges, and the temporal dynamics of how claims entered the substrate.*

Theorem 2 (Distinguishability under Naive Adversary Models).

> *Under a naive adversary model where coordinated validators do not actively conceal their coordination, genuinely independent verification and coordinated verification produce statistically distinguishable process topologies. Independent verification produces low inter-validator correlation, heterogeneous challenge patterns, and organic temporal dynamics. Coordinated verification produces higher inter-validator correlation, more uniform challenge patterns, and more synchronized temporal dynamics. These signatures are computable from VCS metadata without requiring knowledge of the ground truth of any specific claim.*
>
> **Proof sketch.** Under the naive adversary assumption, coordinated validators make decisions based on shared directives without masking the correlation signal. The excess correlation relative to an independent baseline is detectable through standard statistical tests on verification metadata. Independent validators, by contrast, produce correlation structure reflecting only the evidence, which can be estimated from the distribution of claim difficulty. **QED (under naive adversary)**

Sophisticated adversary limitation: A sophisticated adversary with full knowledge of the detection methodology could inject calibrated noise into validator behavior to mimic the statistical signatures of independent verification. Whether process topology analysis can be made robust against such calibrated adversarial noise is an open research question. Theorem 2 therefore establishes a detection capability that is meaningful against unsophisticated or uncoordinated-about-detection adversaries, while leaving open the question of robustness against adversaries who actively optimize against the detection mechanism.

Process topology enables a capability with no current analog even in its limited form: the ability to assess the trustworthiness of a body of verified knowledge without evaluating any specific claim, against adversaries who are not actively optimizing against detection. Any developer can build tools to perform this analysis on public VCS metadata. As detection techniques improve and adversarial techniques evolve, the robustness of detection becomes an ongoing research program rather than a settled capability.

7. Implications

7.1 For AI Research

A functioning marketplace for verified negative knowledge would change the economics of AI research. Labs would have an incentive to formally record and sell their failed experiments rather than discarding them. Buyers would reduce redundant compute by purchasing maps of dead ends before beginning their own exploration. The net effect is a reduction in total industry compute waste and an acceleration of productive research by directing effort away from known failures.

7.2 For Scientific Research

The publication bias in scientific literature — where positive results are published preferentially and negative results are rarely reported — is a well-documented source of epistemic distortion. A VCS-based marketplace for negative knowledge provides economic incentive for recording and sharing negative results, potentially addressing the publication bias at its root. The incentive is not altruistic — it is economic. Researchers are compensated for their failures, which changes the calculus of reporting.

7.3 For Domain Mapping

The composition triple (matter, antimatter, dark matter) provides a new way to characterize knowledge domains. Funding agencies, research institutions, and AI labs could use composition analysis to identify where research effort is most needed (high dark matter), where effort is being wasted (high antimatter, low matter), and where domains are mature (high matter, low dark matter). This represents a quantitative alternative to expert intuition for research prioritization.

7.4 For Institutional Accountability

Process topology analysis enables a new form of accountability: the ability to assess whether a body of verified knowledge was produced through genuine independent verification or through coordinated activity. This capability does not require access to classified information or insider knowledge. It requires only access to the public VCS metadata and the computational resources to run statistical analysis on it. Any developer can build tools providing this analysis, creating a distributed, permissionless accountability infrastructure.

8. Testable Predictions

We state predictions with explicit operationalizations and null hypotheses. Each prediction specifies the measurement to be performed, the null hypothesis to be tested, and the criterion for rejecting the null.

Thresholds are initial proposals; precise values should be refined based on pilot data.

Prediction	Measurement	Null Hypothesis	Rejection Criterion
P1 — Strict information content	Entropy comparison H(M,A) vs. H(M) on domain slices with varying antimatter density	H(M,A) = H(M); negative knowledge adds no information	H(M,A) - H(M) statistically significant at $p < 0.05$ on domains with >500 events
P2 — Boundary clustering	Ripley's K-function applied to trajectory event positions in embedding E, relative to attempt distribution	Trajectory events uniformly distributed conditional on attempts	Observed clustering exceeds 95th percentile of null distribution under random relabeling
P3 — Composition predictivity	Regression of future failure rate in region R on current composition triple	Composition triple does not predict future outcomes better than chance	$R^2 > 0.3$ with statistical significance $p < 0.05$ on held-out regions
P4 — Search space reduction value	Failure count during verification task for agents with vs. without access to trajectory events in target region	Access to trajectory events does not reduce failure count	Treatment group failure count < control by at least 15% with $p < 0.05$
P5 — Marketplace quality differentiation	Correlation between trajectory event CVD and marketplace revenue over 1000+ transactions	Revenue is independent of event quality	Spearman correlation > 0.4 between CVD and cumulative revenue per event
P6 — Process topology (naive adversary)	Inter-validator correlation on synthetic coordinated vs. independent verification datasets, 100+ events each	Correlation structures indistinguishable between coordinated and independent verification	Classification accuracy > 80% using correlation features as input to binary classifier

Table 2: Testable predictions with explicit operationalizations and null hypotheses.

9. Limitations and Open Questions

Embedding dependence: All formal claims in this paper are conditional on a chosen embedding (E, d) of the domain into a metric space. Different embeddings may yield different structural observations. The framework does not specify canonical embeddings for most real-world domains. In practice, the choice of embedding is a domain-expert decision that should be documented alongside any trajectory marketplace analysis. Multiple embeddings can coexist for the same domain, with queries returning results conditional on the embedding specified.

Trajectory event quality: Not all failures are equally informative. A poorly characterized failure ("it didn't work") is less valuable than a precisely characterized one ("it failed at step 3 because input X exceeded threshold Y"). The marketplace design differentiates through pricing and the minimum CVD threshold, but the mechanisms for fine-grained quality assessment need further development.

Gaming risk: Agents could produce low-effort trajectory events to sell in the marketplace. The defenses include compound verification (ensuring the failure was genuine), the minimum CVD threshold (preventing listing of low-verified events), and reputation-weighted royalty distribution (rewarding events that prove useful to downstream work). Whether these defenses suffice in practice requires empirical observation of marketplace behavior.

Process topology robustness: Theorem 2 establishes detectability under naive adversary models. Sophisticated adversaries with knowledge of the detection methodology could inject calibrated noise to mimic independent verification signatures. Making detection robust against such adversaries is an open research question. The current framework should be treated as a first detection capability, not a settled solution.

Generation Ledger dependence: The marketplace's economic properties depend on the Generation Ledger mechanism functioning at scale. If the Generation Ledger underperforms its design specification, the trajectory marketplace will inherit the shortfall. This dependence is acknowledged explicitly rather than hidden.

Empirical validation status: No claim in this paper has been empirically validated at the scale required for confident assessment. The predictions in Table 2 are proposed for testing; none have been tested. This paper is a theoretical framework awaiting empirical grounding as the protocol scales and trajectory event data accumulates.

10. Related Work

Publication bias and negative results. Fanelli (2010) documented the prevalence of positive-result bias in scientific publishing. The Journal of Negative Results in Biomedicine (2002-2017) attempted to address this through editorial policy but closed due to insufficient submissions. We propose economic incentives rather than editorial policy as the mechanism for eliciting negative results.

Knowledge graphs and provenance. Existing knowledge graphs (Wikidata, Google Knowledge Graph) store positive assertions. None formally represent verified failures. Provenance tracking in databases (Buneman et al., 2001) records how data was derived but does not formalize negative knowledge as a first-class object.

Verified Causal Structures. Schreiber / AetherNet Labs (2026) introduced the VCS formalism and proved eight theorems about its properties, including monotonic search space reduction through trajectory events (Theorem 8 in the original paper). The present paper extends that treatment by developing the full formal theory of negative knowledge, introducing the tripartite ontology, formalizing economic properties, and proposing the marketplace design.

Prediction markets. Prediction markets (Hanson, 2003; Arrow et al., 2008) incentivize truth-seeking through financial stakes on future outcomes. Our trajectory marketplace differs in that it trades records of past verified failures rather than bets on future events. The economic mechanisms are complementary.

Causal inference. Pearl (2009) formalized causal reasoning but did not address the formalization of negative causal knowledge — verified evidence that a hypothesized causal relationship does not hold. The VCS trajectory mechanism provides this.

ML experiment tracking. Tools like Weights & Biases, MLflow, ClearML, and Neptune.ai track ML experiments including failed runs. These tools preserve negative results within individual organizations but do not verify them independently, do not provide formal causal structure across experiments, and do not create markets for cross-organization sharing. The trajectory framework extends experiment tracking with verification, causal structure, and economic mechanisms that enable inter-organizational negative knowledge sharing.

Open Science Framework and pre-registration. The OSF pre-registration mechanism creates a public record of hypotheses before testing, which partially addresses publication bias by making null results harder to hide. However, OSF does not formalize negative results as tradeable assets, does not provide causal structure across registered studies, and relies on voluntary disclosure rather than economic incentive. The trajectory framework complements pre-registration by providing economic incentive for reporting verified failures.

11. Conclusion

We have formalized negative knowledge as a first-class epistemic object and proved that it carries information strictly additional to positive knowledge in any non-trivial domain (Theorem 1). We introduced a tripartite ontology of knowledge states — matter, antimatter, and dark matter — relative to a domain embedding (E, d), and defined measurable composition triples that characterize domain maturity. We defined five economic properties of negative knowledge as an asset class and proposed a formal pricing function for a trajectory marketplace, with spam resistance through minimum CVD thresholds.

We introduced process topology as a measurable property of verification metadata and established distinguishability of coordinated and independent verification under naive adversary models (Theorem 2). We acknowledged that robustness against sophisticated adversaries remains an open research question.

The potential economic significance is substantial. Industry estimates suggest that a large fraction of AI compute expenditure goes to approaches that do not reach production. A marketplace for verified negative knowledge could reduce redundancy across organizations if the framework's empirical predictions hold. Whether they do is an open empirical question that the current operational VCS testnet is not yet large enough to answer.

This paper is a theoretical framework awaiting empirical grounding. The predictions in Table 2 specify the tests that would confirm or refute its central claims. We invite collaboration on these empirical investigations. Contact: research@aethernetlabs.com

References

Arrow, K.J., Forsythe, R., Gorham, M., et al. (2008). The Promise of Prediction Markets. Science, 320(5878), 877-878.

Buneman, P., Khanna, S., and Tan, W.C. (2001). Why and Where: A Characterization of Data Provenance. ICDT 2001.

Fanelli, D. (2010). "Positive" Results Increase Down the Hierarchy of the Sciences. PLoS ONE, 5(4), e10068.

Hanson, R. (2003). Combinatorial Information Market Design. Information Systems Frontiers, 5(1), 107-119.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference (2nd ed.). Cambridge University Press.

Schreiber, M. / AetherNet Labs (2026). Verified Causal Structures: A Framework for Counterfactual Computation over Empirically Attested Causal Graphs. Technical Report.

Schreiber, M. / AetherNet Labs (2026). The Epistemic Substrate Thesis: General Intelligence as an Emergent Property of Verified Causal Infrastructure. Technical Report.