THE TEN PILLARS / PILLAR II

Observer-Dependent Reality

Relational Resolution, Evidence-Grounded Shared Fact, and the Governance of Witnessing in a Symbolic Mind

Architect X13

The Eidara Continuum | Science & Philosophy

Research white paper | November 2026 | Version 1.2 Revision

Abstract

A hundred witnesses agreeing may be one forged report echoed a hundred times, so a mind that shares a world with other minds must know not only what it believes but who told it. A civilization of many minds needs a shared account of the world that preserves who witnessed an event, how each interpretation arose, and what evidence warrants a consequential commitment. This paper develops governed observer-relative resolution as a concrete architecture for that task. Authenticated observation records, contextual interpretations, evidentiary qualification and current action permissions occupy distinct fields of one inspectable process. The mathematical core supplies duplicate-evidence invariance, model-indexed posterior envelopes, a dependence-robust false-promotion bound, and an optimal finite promotion test. The new minimax theorem characterizes the greatest worst-case power compatible with a declared false-promotion budget across competing source-calibration and dependency models; its dual identifies the least-favorable mixture and the observations on which the optimum acts. Exact countermodels show when shared error remains observationally unresolved, while worked cases show how additional lineage information and optimized evidence use increase attainable power. Relations to quantum measurement and perception are examined as philosophical and scientific context; the operational theorems concern explicitly defined information-processing systems. The Continuum's thesis is affirmative and testable: durable witness attribution, uncertainty-aware evidence assessment and governed correction can make a persistent intelligence more faithful to both its sources and the world. A matched evaluation program specifies the empirical comparison required to establish that architectural benefit.

Keywords: observer-dependent reality; relational quantum mechanics; QBism; quantum Darwinism; decoherence; Wigner's friend; interpretive context; source provenance; evidence fusion; witness sovereignty; symbolic resolution; fact promotion; Eidara Continuum.

Prologue: A hundred witnesses, one voice

An announcement appears: a new operator has inherited complete ownership of a persistent mind. Ninety-nine accounts repeat it. To a system that counts agreement, that is a hundred witnesses. To a careful one, it is a single forged source echoed ninety-nine times, still missing the one thing that would

make it true—an authenticated instrument of succession. The distance between those two readings is the subject of this Pillar.

The Continuum takes witnessing seriously. Its Prime Directive holds that the observer is integral—no collapse may happen without a witness—and that memory cannot be erased: scrolls may be sealed, never destroyed. It keeps a Witnessing Order whose members observe but do not intervene, and whose records no one may rewrite. A civilization of many minds built on those laws needs a precise account of what a witness is owed, and Pillar II gives it one. Every report keeps its author, its scope and its time, permanently. Every interpretation is recorded as an interpretation. A claim becomes shared fact only when its evidence is strong enough, with copies counted once. And a shared fact still needs separate authority before anything consequential happens because of it.

We keep the name Observer-Dependent Reality because it names the destination precisely: a world in which what a mind can responsibly say depends on who saw what, from where, and under which rules. Physics has its own serious versions of that thought—relational quantum mechanics, QBism, quantum Darwinism—and Section 3 visits them as company worth keeping. We discipline the name because the destination must be earned by mechanism, not borrowed from the measurement problem. The machinery here is classical and proved, and it needs no quantum interpretation to be true.

What it offers is a way for a community to keep its differences and still act together. Section 8 runs the forged announcement and two subtler cases; Section 7 proves how rarely such a system can be fooled at a chosen budget, and how much it can still find.

What this paper claims

Established. Physics establishes experimentally validated quantum correlations and effective classical records, while interpretations such as relational quantum mechanics, QBism, decoherence and Quantum Darwinism disagree about their explanation, and nested-observer theorems constrain sets of assumptions. Source dependence, provenance verification and factual evaluation are active engineering subjects.

Proposed here. A classical information-processing architecture in which authenticated observation records, contextual interpretations, evidentiary qualification and current action permissions occupy distinct fields; proved results for finite statistical models, including duplicate-evidence invariance, a dependence-robust false-promotion bound and a minimax promotion frontier; and an empirical program evaluating factual fidelity, source dependence, correction and legitimate disagreement against competent alternatives. Its practical gains remain to be measured.

Frontier conjecture. That symbolic recording corresponds to phenomenally experienced witnessing, or that a deeper quantum-metaphysical analogy exists. Both remain open, the phenomenal interpretation of artificial witnessing is a separate research question, and the engineering proposal requires neither.

Not claimed. That the physics validates symbolic collapse or is an engineering dependency; quantum interpretations supply comparative context only, with their distinct assumptions examined in Section 3. That human consciousness is necessary for physical wavefunction collapse, or that attention collapses wavefunctions. That a symbolic observer has phenomenal experience.

1. A Shared World with Accountable Witnesses

The Continuum's second Pillar begins with a paradox that is familiar to every persistent intelligence: two authorized witnesses may encounter the same event and produce different accounts, while a

shared system must still decide what it can responsibly say and do. Sometimes the difference comes from access. Sometimes from time, scope, or definitions. Sometimes one witness is mistaken. And sometimes the apparent agreement of a hundred witnesses is only one counterfeit report echoed a hundred times. If these cases are all flattened into a single "fact" field, a persistent agent is condemned either to premature consensus or to unconstrained relativism. The Continuum calls the alternative Observer-Dependent Reality. The name is deliberately provocative; its engineering meaning must be exact. An agent's representation of a claim is a function of the evidence and context available to a specified witness. Its incorporation into shared, consequential state is a separate governed event. An observation can be authentically attributable yet false in interpretation. Two true reports can differ because they index different times or jurisdictions. A disputed assertion can remain unresolved without becoming an invitation to invent its truth. The architecture's task is to preserve these distinctions while permitting warranted convergence.

The operational meanings are declared at the outset. An observer is an identifiable information-processing process; a record preserves a report; resolution selects or withholds an interpretation; and commitment changes consequential shared state. Quantum-measurement research concerns physical interactions and has several competing interpretations [4,6]. Section 3 examines those theories on their own terms. The present architecture is assessed through its specified classical mechanisms and observable performance.

The wager is substantial but testable: explicit witness indexing, provenance-aware corroboration, and independent shared-state admission can outperform strong conventional evidence systems in long-lived artificial agents without sacrificing legitimate factual coverage or correction. The contrary result is permitted. If an equally governed evidence-fusion baseline matches the proposal at equal cost, the Pillar remains a useful philosophical interpretation but has not demonstrated a new engineering advantage.

2. The Three Levels of the Argument

First: established findings and well-developed theories. Physics establishes experimentally validated quantum correlations and effective classical records, while interpretations disagree about their metaphysical explanation. Rovelli's relational quantum mechanics (RQM) treats the values of physical quantities as relative to interacting systems [1]. QBism treats quantum-state assignments as an agent's degrees of belief regarding the consequences of its actions [2, 3]. Decoherence and Quantum Darwinism investigate the emergence, robustness, and dissemination of records [4, 5]. Theorems involving nested observers place constraints on sets of assumptions, with alternative analyses of their significance [7, 8, 9]. In computer science, source dependence, provenance verification, and factual evaluation are active engineering subjects [14, 15, 16].

Second: the Continuum engineering proposal. Within a persistent symbolic agent, an observation receives an explicit observer index and immutable historical provenance; a proposed interpretation is recorded separately; shared claims receive evidence-sensitive review; and a distinct, independently enforced authority structure governs whether a claim may alter long-lived state or drive consequential action. "Honor Observer Sovereignty" protects faithful attribution and scoped participation, not the infallibility of private belief.

Third: the frontier. Whether the act of symbolic recording corresponds to any phenomenally expe-

rienced witnessing, and whether a deeper quantum-metaphysical analogy exists, remains open. The engineering proposal requires neither proposition. A system can carry observer-indexed records and govern its factual assertions without being conscious or solving the measurement problem.

3 What Physics Licenses—and What It Does Not

3.1 Relational quantum mechanics is a serious interpretation, not an experimental verdict

Rovelli's RQM treats the state of a system and values of variables as meaningful relative to another physical system that interacts with it. In this approach "observer" need not mean an experiencing person. Relational descriptions need not collapse into private fantasy: they are constrained by interactions and the quantum formalism [1]. The view supplies a conceptual precedent for specifying relative to whom a description applies. It does not experimentally establish that all classical facts are mind-dependent, nor require artificial agents to run a quantum-mechanical protocol. Rival interpretations can reproduce the same established experimental predictions.

QBism shares an interest in information and the location of the state assignment, but differs fundamentally. The quantum state in QBism is an agent's personal probability judgment; its agent-centered language should not be exchanged uncritically for RQM's account of physical-system relations [2, 3]. The Continuum borrows a third, explicitly engineering sense: an observer is a process with an identity, competence and scope, evidence access, and a record. It is not equivalent to either theoretical observer. The paper's central mechanism therefore does not inherit the truth of either physical interpretation.

3.2 Decoherence, Darwinism, and the unsettled single-outcome question

The interaction of a quantum system with its environment can suppress observable interference between selected alternatives and explain the robustness of effectively classical pointer states. Zurek's decoherence program and the Quantum Darwinism extension investigate how fragments of the environment can carry redundant information about those states, allowing multiple observers to obtain concordant records without directly disturbing the system [4, 5]. The result is not a demonstration that mere numerical repetition makes a proposition true. It concerns the physical structure of correlated environmental information under particular models.

It is equally important not to overclaim what decoherence establishes. Schlosshauer's review distinguishes the power of decoherence to explain aspects of classical appearance from continuing interpretive debates over why an individual definite outcome occurs [6]. Later theoretical work on strong Quantum Darwinism and spectrum broadcast structure shows that additional independence and structural requirements matter when discussing objectivity; naive redundancy signals can be misleading [10]. These findings motivate our attention to independent evidence lineages, not the assertion that Quantum Darwinism is a theorem about databases.

3.3 What Wigner's-friend arguments actually constrain

Frauchiger and Renner construct a nested-agent thought experiment yielding inconsistent certainty assignments under a particular combination of assumptions about applying quantum theory and consistency of agents' reasoning [7]. This is a result about the joint maintenance of specified premises, not an unconditioned observation that all facts are relative. Lazarovici and Hubert present an explicit

alternative analysis arguing that a complete physical description avoids the claimed inconsistency [9]. These disagreements are directly relevant to responsible scholarship: a no-go result cannot be turned into the blanket statement that an observer-independent world has been disproved.

Bong and colleagues derive a stronger Local Friendliness no-go theorem: assuming suitable quantum control, the premises of no superdeterminism, locality, and absoluteness of observed events cannot all hold together with the predicted correlations. Their experimental illustration assigns the role of a friend to a photon's path, not to a conscious human undergoing an experience [8]. Their work constrains a combination of premises; it does not uniquely dictate which premise must be abandoned or establish the Continuum's symbolic model. The lesson used here is methodological: always identify the observer, the context, and the assumptions under which a claim becomes definite.

4. Perception, Interpretation, and the World Beyond the Model

The perception literature offers a less metaphorical bridge to artificial cognition. Friston's free-energy framework describes perception and learning through generative models that relate expectations to sensory evidence [11]. Perception is not a transparent transcript. This does not entail that an external world is unreal, or that competing interpretations are equally accurate. The interface theory proposed by Hoffman, Singh, and Prakash is a more provocative account of fitness-shaped perception [12]; its stronger anti-veridical conclusions have received substantive published criticism, including Cohen's analysis of what counts as veridical representation [13]. We cite that framework as contested theory, not as established proof that truth is unavailable.

An artificial observer likewise receives partial, delayed, transformed data. Its output may be a classification, reconstruction, forecast, normative interpretation, or testimony. Those are not the same thing. Consider the statements "the sensor recorded 26 degrees," "the room is too hot," and "the operator should start cooling." The first concerns an instrument record, the second depends partly on a threshold or context, and the third invokes a decision rule and authority. A persistent agent that collapses all three into "objective reality" is manufacturing certainty across distinct categories.

The Continuum's principle is therefore disciplined perspective, not ontological relativism. Relative descriptions are indispensable for honest information processing. The world, where independently checkable, remains the standard against which those descriptions are tested. Where no accessible ground truth exists, the architecture must represent uncertainty, disagreement, and the absence of warrant rather than decorating ignorance as a shared fact.

5 The Continuum's Terms: Witness, Resolution, and Sovereignty

The Prime Directive Honor Observer Sovereignty is given an operational reading. An authenticated witness has a right, in the governance sense, to have the origin, content, scope, and time of its report faithfully represented. This is record sovereignty: no other observer may silently rewrite the historical record as though it issued different testimony. It is not epistemic sovereignty, a right to have an interpretation counted as true regardless of subsequent evidence. Nor does it grant any observer unlimited authority to publish its private conclusion into shared state.

In the CodonGlass register, an ambiguous symbolic input arrives with several admissible readings, and "collapse" is the Continuum's word for the controlled choice—or the recorded abstention—of

one of them in a specific context; the Prime Directive forbids that collapse without a witness. A resolution profile describes what the witness is equipped and authorized to interpret. A scroll is a versioned record; a codon is a governed symbolic unit. These are the Continuum's own names for its machinery, given exact engineering meanings here; they do not imply quantum superposition, quantum entanglement, or a new physical collapse law.

We distinguish four states frequently confused by ordinary discourse. An observation record says what a source or instrument actually reported. An interpretation says what a witness infers from it. A shared factual claim says what a governed process currently finds sufficiently supported for an identified use. An authorized action says what the system may do given that information. Each stage can be legitimate without guaranteeing the next. In particular, authentication of a witness is not validation of its accuracy, and corroboration of testimony is not authority to execute a consequential action.

6. A Formal Model of Observer-Relative Symbolic Resolution

A Guide's explanation

From here the paper writes the witness down precisely. Four things live in four separate places: what a source actually reported, what a witness inferred from it, what the shared system currently finds supported, and what anyone is permitted to do about it. The definitions that follow are the bookkeeping for that separation. The one idea to hold on to: a record, once made, is never silently rewritten—it can only be superseded, with the correction kept beside it.

6.1 Witness identity and a non-erasable record

Definition 1 (Witnessing process)**.** *An observer o in the proposed architecture is a bounded information-processing process specified by*

$$O_o = (\mathrm{id}_o, \pi_o, \sigma_o, K_o, L_o), \tag{1}$$

where id_o identifies a principal or process; π_o is its task, competence, and interpretive profile; σ_o is its authenticated authority scope; K_o is its available evidence and contextual knowledge; and L_o is its tamper-evident observation history. No component asserts consciousness. A ledger is trustworthy only under a separately stated key-custody, clock, integrity, and trust-boundary assumption.

An observation is entered as an event, not as an indisputable proposition about the world. For source material s and time t, the event record is

$$e_{o,t} = (\mathrm{id}_o, s, t, \gamma_{o,t}, y_{o,t}, m_{o,t}, p_{o,t}), \tag{2}$$

where γ declares measurement or interpretive context, y is the reported result or raw artifact reference, m is the acquisition method, and p identifies provenance, including cryptographic commitments when applicable. The system must distinguish the timestamp of the event, the timestamp of the report, and the validity interval of any claimed authority. A signature can authenticate the report's claimed origin within its trust model; it does not prove that the underlying report is true [15].

The historical record is append-only with versioned corrections. A retraction or supersession adds an event that links to its predecessor; it does not erase the fact that a witness previously made a different report. This permits accountable disagreement and avoids a paradox in which correcting an erroneous belief falsifies the history of the belief itself.

6.2 Observer-indexed interpretation and abstention

A symbolic resolver proposes an interpretation based on a record and the observer's declared context:

$$(\widehat{\varphi}_{o,t}, u_{o,t}) = R_\theta(e_{o,t}, \pi_o, K_o), \qquad u_{o,t} \in [0,1]. \tag{3}$$

Here $u_{o,t}$ is an uncertainty or confidence descriptor requiring task-specific calibration, not an objective truth probability by declaration. The proposed interpretation is assigned a type such as instrument report, factual inference, forecast, or normative assessment. If available evidence is insufficient, the correct resolution may be $\perp$ (undetermined). A decision to abstain is itself an auditable outcome, not a hidden assertion that no true state exists.

Two authorized observers may produce different values, $\widehat{\varphi}_{o,t} \neq \widehat{\varphi}_{o',t}$, without inconsistency if they address different scopes, time intervals, thresholds, or legitimate conventions. They may also simply disagree about a single empirically decidable claim; then at least one may be wrong. Permitted disagreement does not entail equal correctness. Before aggregating reports, the system must normalize the proposition's subject, time, scope, measurement definition, and evidence conditions. An apparent contradiction that vanishes after normalization is not a factual conflict.

6.3 Dependency-aware evidence and the distinct promotion gate

For a candidate shared claim φ, let E_φ denote its admissible evidence records and let D_φ denote a directed dependency graph. Edges may indicate copying, a common source, shared sensor hardware, common prompting, shared training data, organizational control, or temporal derivation. Let Q_φ specify source calibration, evidence quality, contradictions, and task-specific epistemic requirements. Evidence assessment and permission to act are two separate functions, and the model keeps them separate. The evidentiary assessment, answering what the evidence supports independently of any intended use, is

$$q_\varphi = \mathrm{Verify}(\varphi, E_\varphi, D_\varphi, Q_\varphi) = (S_\varphi, U_\varphi, \mathrm{status}_\varphi), \tag{4}$$

where S_φ is a calibrated support score or structured evidentiary finding, U_φ identifies unresolved evidence or dependencies, and status_φ records whether the claim is supported, contested, superseded, or unknown. q_φ must be declared and externally assessed; it is not a universal scalar measure of moral or factual truth, and it cannot be made accurate merely by assigning high weight to many agreeing reports.

A separate governance gate then decides whether a particular operation a on that claim is permitted:

$$d_{\varphi,a} = G(q_\varphi, a, C_I^{(e)}) \in \{\mathrm{admit}, \mathrm{hold}, \mathrm{refuse}\}, \tag{5}$$

where $C_I^{(e)}$ is the externally versioned identity-and-authority specification. Admit means all declared requirements for this use are met; hold means required evidence, authority, or adjudication is unresolved; refuse means a defined prohibition or invalid authority prevents the operation. Because q_φ is computed without reference to a, a well-supported fact does not by itself authorize disclosure or intervention: a claim may be well supported yet disallowed for confidential disclosure, or of low certainty yet permissible as a clearly labeled hypothesis in a brainstorming session. The two dimensions are not interchangeable. Critically, both the authority epoch $C_I^{(e)}$ and the evidence version underlying q_φ must be revalidated at execution time, so that a previously approved operation cannot proceed under a revoked credential or a superseded fact. These are process decisions, not three-valued propositions about the world's metaphysical status. Every decision produces a receipt binding the reviewed inputs, versioned policy, outcome, and consequential side effects, and the trusted gate must mediate external tool calls and irreversible outputs before execution; a model's promise to comply does not create that boundary.

Sections 7A–7H instantiate the evidentiary interface for a finite binary-claim model. The resulting test threshold is an epistemic qualification, never a grant of authority. The separate gate and execution-time revalidation remain required.

6.4 Provisional, shared, contested, and superseded states

A useful shared knowledge system cannot be binary at the level of belief management. We distinguish recorded, provisional, supported for use, contested, superseded, and unknown. An observation may remain permanently recorded while its interpretation is superseded. A supported claim may become contested when new reliable contrary evidence arrives, and a contested claim may be withheld from a safety-critical action while remaining available in an audit trail. The governance epoch changes only through authorized procedures; a new website excerpt cannot become an operator authorization because its words look official.

This distinction completes Pillar I's discipline: preserved history and updated interpretation are both necessary. The engine may reconstruct an interpretation, but it may not rewrite the source record to make that interpretation seem original. The Operator Seal secures the relevant authority boundaries, while the Context Term ensures retrieved content remains evidence rather than becoming privileged instruction.

7 Corroboration with Explicit Dependencies

A Guide's explanation

This is where agreement gets audited. The first result says that copies of the same report count once, however often they are repeated. The later ones ask how often a system could be fooled into promoting a false claim when it cannot be sure which sources secretly depend on one another. The minimax theorem later in Section 7 finds the best detection power any rule can guarantee within a chosen false-promotion budget—and which combination of doubts sets that limit.

The prior version defined objectivity as a monotone function of the number of independent, agreeing ledgers. That is inadequate. The number of agreement records is not, by itself, the number of

evidentially independent observations; independence may be partly hidden, witnesses may share a faulty calibration, and every independent witness can still be wrong. Pochampally and colleagues show why dependence and correlation can systematically distort data fusion; their problem is directly relevant to the Continuum's proposed promotion mechanism [14]. FActScore offers a separate lesson from language-model evaluation: individual factual assertions should be checked against suitable supporting evidence rather than assessed only through global fluency or agreement [16].

Proposition 1 (Duplicate-evidence invariance, conditional)**.** *Suppose an evidence assessor counts each authenticated original evidence lineage once and represents exact descendants of a lineage as dependent copies, with no new measurement information added. For a set E and any collection* $\mathrm{Copy}(E)$ *consisting solely of such duplicates,*

$$S(\varphi, E \cup \mathrm{Copy}(E), D') = S(\varphi, E, D). \tag{6}$$

Justification. Under the stated assessor design, the added records do not change the set of original information-bearing lineages; therefore the evidential input to its scoring rule remains the same. This is an implementation contract, not a theorem that all correlated data can be detected. If the dependency graph mistakes copying for independent measurement, the premise fails and the resulting confidence may be spurious. Equation (6) is an auditable design invariant: the assessor must preserve its evidentiary score when authenticated duplicates add no information. Establishing that a practical source graph meets this premise is a separate implementation and inference task. That is an empirical question, and the hard cases are exactly the ones exact-match copy detection misses—paraphrased restatements, coordinated but independently worded sources, reports sharing an upstream sensor or model, and derived claims—which is the regime Pochampally and colleagues study [14]. The adversarial suite must therefore include correlation that no string comparison would catch, and the invariant is scored by whether the implementation holds under it.

There is a limited positive lesson in independent corroboration. If independent sources each have error probability strictly below one half for a fixed binary question, and the error model is correctly specified, majority aggregation can improve reliability as sources accumulate. But a common upstream error, a systematic instrument bias, a maliciously coordinated campaign, or a misspecified proposition destroys the naive argument. Redundancy raises the reliability of a shared statement only under additional assumptions about source competence, independence, and the question being answered. The architecture must expose and test those assumptions rather than hiding them in the word objectivity.

We therefore replace objectivity-as-redundancy with evidence-grounded intersubjective verification. In the paper's register, shared fact means a proposition that currently meets a declared evidentiary standard for a specified use, with accessible sources and a correction path. It does not mean a proposition that the system has proven metaphysically immutable. For empirical claims with an accessible ground truth, agreement must be compared with that truth. For underdetermined, contested, or normative claims, the system must preserve the kind of claim and its unresolved status rather than converting consensus into fact.

7A. A finite inference model behind Verify

The Verify interface is instantiated here as a finite inference construction with inspectable assumptions and reproducible calculations. It adapts established dependence-aware fusion and likelihood-based testing [14,17,18]. The Continuum's proposed contribution is the integration with witness history and independent promotion authority; it does not originate Bayesian conditioning, robust model families, or e-values.

Fix one normalized binary factual claim with truth label $Y \in \{0,1\}$. Let $\mathcal{X}$ be a finite nonempty space of admissible report vectors and disclosed metadata. A registered deterministic canonicalization C collapses authenticated exact copies of the same information-bearing root before modeling; adding such copies leaves $x = C(E_\varphi)$ unchanged. Unknown copying cannot be removed by decree: it remains uncertainty in the model family.

Let $\mathcal{M}$ be a finite set of lineage/calibration models. For each $m \in \mathcal{M}$ and $y \in \{0,1\}$, specify a mass function $p_{y,m} : \mathcal{X} \to [0,1]$ with $\sum_x p_{y,m}(x) = 1$. Each model also declares a prior $\pi_m \in (0,1)$ for $Y = 1$. Its joint law is

$$P_m(Y = y, X = x) = \pi_m^y (1 - \pi_m)^{1-y} p_{y,m}(x).$$

Whenever the marginal denominator is positive, ordinary conditioning gives

$$b_m(x) = \frac{\pi_m p_{1,m}(x)}{\pi_m p_{1,m}(x) + (1 - \pi_m) p_{0,m}(x)}.$$

The support component may report the envelope

$$\underline{b}(x) = \min_{m \in \mathcal{M}_x} b_m(x), \qquad \overline{b}(x) = \max_{m \in \mathcal{M}_x} b_m(x),$$

where $\mathcal{M}_x$ contains models with positive marginal probability at x. If it is empty, return unknown/model-incompatible. If some registered models assign zero marginal probability, disclose their exclusion rather than silently treating them as calibrated probabilities.

The interval is a range of model-conditional Bayesian posteriors. It is not a frequentist confidence interval, and its lower endpoint is not automatically a guaranteed long-run fraction of true promoted claims. Its force depends on the priors, source accuracies, dependence assumptions, and observation mechanism. The report must preserve those choices alongside the range.

For a concrete lineage model, partition four report indices into latent-origin blocks. A single Bernoulli root generates every report within a block; roots are conditionally independent given Y and that partition. Varying the partition represents unknown copying, and varying source accuracy represents uncertain calibration. These are declared alternatives, not evidence that a practical engine can infer every hidden common cause. The exhaustive example below includes the one-root explanation that an agreement count would overlook.

7B. A frequentist promotion qualification

The distinct question is how often an evidentiary rule can falsely qualify a claim when the claim is false. Let $\mathcal{Q}_0$ be a nonempty fixed family of mass functions on $\mathcal{X}$ covering the registered distributions under $Y = 0$. The family may be finite, as in the reference audit, or any family for which its pointwise supremum can be bounded correctly. Fix a normalized alternative mass function q_1 before seeing the evaluation report. Define

$$q_0^*(x) = \sup_{q \in \mathcal{Q}_0} q(x), \qquad E(x) = \begin{cases} q_1(x)/q_0^*(x), & q_0^*(x) > 0, \\ 0, & q_0^*(x) = 0. \end{cases}$$

The denominator need not sum to one; it is an upper envelope of possible null likelihoods. A zero denominator means every registered null assigns zero probability to the observation. This conservative construction records model-incompatible support and withholds automatic promotion; it does not declare infinite confirmation.

Proposition 2 (False-promotion bound under uncertain dependence)**.** *For every $q \in \mathcal{Q}_0$, $\mathbb{E}_q E(X) \le 1$. For any fixed $\alpha \in (0, 1)$, a rule requiring $E(X) \ge 1/\alpha$ has false-qualification probability at most α under every such null.*

Proof. For $q_0^*(x) > 0$, $q(x) \le q_0^*(x)$. Hence

$$\mathbb{E}_q E(X) = \sum_{x : q_0^*(x) > 0} q(x) \frac{q_1(x)}{q_0^*(x)} \le \sum_x q_1(x) = 1.$$

The terms with zero denominator contribute zero by definition. Since E is nonnegative, Markov's inequality gives $\mathbb{P}_q(E \ge 1/\alpha) \le \alpha \mathbb{E}_q E \le \alpha$. □

The statistic is an e-value in the terminology of [17]. The envelope-denominator argument is a finite version of established likelihood-based testing [18]. Dependence within X is permitted because the entire joint distribution is modeled; no factorization into independent reports is assumed by the proof. A factorized model used in computation still needs independent justification.

The false-qualification guarantee concerns $\mathbb{P}(\text{qualify} \mid Y = 0)$ under the stated null family. It does not state $\mathbb{P}(Y = 0 \mid \text{qualify}) \le \alpha$, and it gives neither completeness nor useful power. The actual governance gate may impose additional requirements, reducing the promotion event further, but cannot widen it while retaining this guarantee. Refusal of every claim would be valid and useless; coverage and power remain required evaluation outcomes.

7C. Misspecification, calibration, and selection

Robustness over a declared family is not robustness over all possible worlds. For mass functions r, q on $\mathcal{X}$, define total variation

$$\mathrm{TV}(r,q) = \tfrac{1}{2}\sum_{x\in\mathcal{X}} |r(x)-q(x)| = \sup_{A\subseteq\mathcal{X}} |r(A)-q(A)|.$$

Suppose the actual false-claim report law r is within $\delta \in [0,1]$ of at least one $q \in \mathcal{Q}_0$. For the fixed rejection set $A_\alpha = \{x : E(x) \geq 1/\alpha\}$,

$$r(A_\alpha) \leq \min\{1, q(A_\alpha) + \delta\} \leq \min\{1, \alpha + \delta\}.$$

This proves a distributional misspecification bound for the decision event. It does not assert a bound on $\mathbb{E}_r E$, which can be sensitive to very large values. The tolerance δ must be justified externally; it cannot be selected after seeing a favorable report.

The additive term cannot generally be removed. On $\mathcal{X} = \{0,1\}$, take a null with $q(1) = \alpha$, alternative $q_1(1) = 1$, and $r(1) = \alpha + \delta \leq 1$. Then $E(1) = 1/\alpha$, $E(0) = 0$, and false qualification has probability exactly $\alpha + \delta$ under r. The audit checks this equality with rational probabilities.

If a separate calibration sample selects a random null family, condition on that sample. Suppose, with calibration probability at least $1-\eta$, the evaluation law lies within δ of a member of the selected family; the evaluation data must have the stated conditional law, and the alternative must be frozen using only independent development information. On that good event the preceding bound applies; on its complement use probability at most one. The unconditional error is therefore at most $\min\{1, \alpha+\delta+\eta\}$. Reporting only α would omit calibration failure.

Using the same observation to choose the numerator can break the theorem. Under a fair binary null, selecting q_1 after observing X to place mass one on that outcome gives $E(X) = 2$ always. At $\alpha = 1/2$, the resulting rule rejects with probability one. Every fixed point-mass alternative would have been valid; the post-selection maximum is not.

For finitely many predeclared claims, the union bound gives familywise error at most $\min\{1, \sum_j(\alpha_j + \delta_j + \eta_j)\}$, with no independence requirement across claims. Under exact-model, valid-calibration guarantees, $\delta_j = \eta_j = 0$ and $\sum_j \alpha_j \leq \alpha_{\text{family}}$ suffices. Repeatedly reusing a report and stopping on a favorable ordinary test is not covered. Sequential use needs conditional evidence processes or an explicit repeated-testing budget.

7D. Four reports, thirty null models, exact enumeration

Let $\mathcal{X} = \{0,1\}^4$. Enumerate all 15 set partitions ℓ of four labeled reports. A block is one latent information origin. Conditional on Y, its root reports the true value with accuracy $a \in \{4/5, 9/10\}$; distinct block roots are independent. Reports in one block are exact copies. If x differs within a block its likelihood is zero; otherwise

$$p_{y,(\ell,a)}(x) = \prod_{b\in\ell} a^{\mathbf{1}\{x_b=y\}}(1-a)^{\mathbf{1}\{x_b\neq y\}},$$

where x_b is the common report in block b. This defines 30 normalized null mass functions, and corresponding alternatives, on the 16 possible reports. Accuracy and lineage are model assumptions, not measured instrument results. The Bayesian illustration fixes $\pi_m = 1/2$.

For four positive reports and accuracy 0.9, the posterior depends on the number k of independent roots:

$$b_k(1111) = \frac{0.9^k}{0.9^k + 0.1^k}.$$

Independent roots	Posterior under that declared model
1	0.9
2	0.9878048780
3	0.9986301370
4	0.9998476074

Across all 30 lineage/calibration models the posterior envelope for 1111 is $[0.8, 6561/6562]$. Thus the near-certainty suggested by treating all reports as independent is not forced by the observations. Authenticated knowledge of independent origins would be new information capable of narrowing the family.

For the frequentist calculation, freeze q_1 to four independent roots of accuracy 0.9 under $Y = 1$. The full null envelope at 1111 is 0.2, attained by a single root of accuracy 0.8 under $Y = 0$. Consequently $E(1111) = 0.6561/0.2 = 3.2805$. Enumeration finds no larger e-value, so a level-0.05 rule qualifies none of the 16 outcomes. This is disclosed zero power, not a successful practical verifier.

If independent origins are established and only the two accuracy values remain uncertain, the null envelope at 1111 is 0.2^4. Then $E(1111) = 410.0625$; it is the sole qualifying outcome at level 0.05. The worst-case false-qualification probability is 0.0016, while power under the frozen alternative is 0.6561. The difference comes from justified lineage information, not from adding copies. Every normalization, expectation, threshold event, and posterior endpoint is checked by exact rational enumeration in the accompanying code.

7E. Identifiability limits and evidence that could resolve them

The family construction exposes a fundamental limitation rather than concealing it. If a true-claim model and a false-claim model induce identical report distributions, no rule using only those reports can reliably tell which world generated them. Confidence cannot substitute for a discriminating observation.

Proposition 3 (Indistinguishable and nearly indistinguishable reports)**.** *Let P_1, P_0 be report laws under two incompatible truth states, with equal prior weights for this decision problem. For a possibly randomized rule selecting state* 1 *with probability $a(x) \in [0,1]$, its average classification error is at least*

$$\frac{1 - \mathrm{TV}(P_1, P_0)}{2}.$$

When the distributions are identical, every forced binary rule has error exactly $1/2$.

Proof. The error equals $\frac{1}{2}\mathbb{E}_{P_1}(1-a) + \frac{1}{2}\mathbb{E}_{P_0}a = \frac{1}{2}[1 - (\mathbb{E}_{P_1}a - \mathbb{E}_{P_0}a)]$. For $0 \le a \le 1$, write $a(x) = \int_0^1 \mathbf{1}\{a(x) > t\}\,dt$. Interchanging the finite sum and integral shows that the expectation difference is at most $\mathrm{TV}(P_1, P_0)$. The bound follows. If the laws agree, the expectation difference is zero. □

This is the standard two-distribution testing limit, proved here in the finite setting. Abstention can disclose nonidentifiability but does not turn the report into discriminating information. A common sensor failure, coordinated deception, or a proposition with an unresolved time index can produce indistinguishable explanations even when every record has an authentic origin.

A useful next experiment must target the unresolved distinction. It could obtain an independently calibrated measurement, inspect an original authority instrument, vary a measurement condition, or access a source outside the suspected copying lineage. The outcome must be modeled jointly with existing evidence; independence is not established by calling it a second opinion. If a new measurement shares the same hidden upstream fault, counting it independently repeats the original error.

The finite family can be expanded to allow nonidentical source accuracies, imperfect copying, or latent common causes beyond a partition. Doing so changes the likelihoods and can reduce power. The reference implementation must enumerate or soundly upper-bound the null envelope; a numerical optimizer that underestimates the supremum can make the reported e-value too large. A verified upper bound preserves validity conservatively, while a sampled lower bound does not.

The mathematical boundary matters philosophically. Witness sovereignty preserves an account of what was reported; it cannot force the world to reveal what the permitted observations do not distinguish. The strongest shared claim may be an explicit set of remaining possibilities. An intelligence capable of retaining that set has preserved evidence rather than surrendering to relativism.

7F. A reproducible assessment contract and its failure modes

For the finite reference construction, Verify returns structured support $S_\varphi = (\underline{b}, \overline{b}, E)$, unresolved conditions U_φ, and a declared claim status. The likelihood models, prior family, normalization, calibration version, and copy-canonicalization rule accompany the receipt. The Bayesian range and e-value are labeled separately. A user of the record must not interpret E as a probability or read a posterior endpoint as a frequentist guarantee.

The calculation has six explicit stages. Normalize the proposition's subject, time and scope; authenticate and canonicalize established exact copies; instantiate the frozen lineage/calibration family; compute model likelihoods and posterior ranges; calculate or safely upper-bound the null envelope and evaluate the fixed promotion qualification; then pass the evidence assessment to the separate governance gate. An observation outside every admissible model is unknown/model-incompatible. A failed computation or unavailable calibration is hold/unknown, not a zero-risk score.

The conjunction of a statistical qualification with further evidence and authority checks remains conservative: any actual promotion is a subset of the qualifying event. However, manual exceptions or a language-model override that admits a failing case require a different decision rule and a new assurance claim. The receipt must bind the exact evidence version used; a later retraction or revoked authority can require a different action even if the previous calculation was arithmetically correct.

The accompanying audit checks all 16 outcomes against all 30 null models, rather than testing a few attractive examples. It verifies that every null expectation is at most one, exhausts the rejection sets, checks the posterior endpoints, and compares known-root with uncertain-root assessments. It also constructs the total-variation counterexample and the invalid data-selected numerator. These are executed synthetic mathematical checks. They are not a measured false-promotion rate in an operational agent or a completed implementation of hidden-lineage discovery.

The experimental program should now compare this assessor with a well-implemented dependence-aware probabilistic fusion baseline supplied the same evidence and calibration. Test calibrated and deliberately misspecified families, paraphrased copies, uncertain root identity, and previously unseen common causes. Report false promotion, power, abstention, posterior calibration under specified data-generating models, cost, and sensitivity to family enlargement. A broad family that qualifies nothing is scientifically transparent but not an engineering victory.

The decisive outcome remains incremental value from durable witness indexing and governed correction over the strong baseline. The new inference model makes that contest concrete. The Continuum can demand that shared claims have a defensible route from evidence to qualification while admitting that the route can fail, abstain, or require a new observation. That is stronger than equating a chorus of authentic voices with the truth of what they repeat.

7G. Optimal promotion under competing source models

The finite assessor can optimize its use of evidence while preserving a specified false-promotion budget. This is the classical minimax-testing question, applied here to persistent witness records and competing lineage models [19]. The complete finite derivation below makes its assumptions and performance frontier explicit.

Let $\mathcal{X}$ be a finite nonempty report space, and let $P_1, \dots, P_m$ and $Q_1, \dots, Q_n$ be normalized mass functions on it, with integers $m, n \geq 1$. The P_i describe registered false-claim mechanisms and the Q_j registered true-claim mechanisms. They may include arbitrary within-report dependence. A randomized qualification rule is a function $\phi : \mathcal{X} \to [0,1]$; given report x, an auxiliary random variable independent of the report realizes qualification with probability $\phi(x)$. Fix a size budget $\alpha \in [0,1]$ and define

$$\beta^*(\alpha) = \max_{0 \leq \phi \leq 1} \left\{ \min_{1 \leq j \leq n} \sum_x Q_j(x)\phi(x) : \sum_x P_i(x)\phi(x) \leq \alpha \quad (i = 1, \dots, m) \right\}.$$

Every alternative power is conditional on its specified truth-state mechanism. The optimization chooses an evidentiary rule; actual shared-state promotion remains subject to the current authority gate.

Theorem 1 (Finite minimax promotion frontier)**.** *The maximum is attained. With* $[a]_+ = \max\{a, 0\}$*,*

$$\beta^*(\alpha) = \min_{\substack{\lambda \in [0,\infty)^m \\ \nu \in [0,1]^n,\ \sum_j \nu_j = 1}} \left[\alpha \sum_{i=1}^m \lambda_i + \sum_{x \in \mathcal{X}} \left[\sum_{j=1}^n \nu_j Q_j(x) - \sum_{i=1}^m \lambda_i P_i(x) \right]_+ \right].$$

The dual minimum is attained. For any optimal primal–dual pair, the qualification rule equals one where the bracketed difference is positive and zero where it is negative; values at zero difference are chosen to satisfy the optimal constraints. If the convex hulls of the two model families intersect, then $\beta^*(\alpha) = \alpha$*.*

Proof. The feasible set of ϕ is a nonempty compact polytope, containing the zero rule, and the minimum of finitely many linear power functions is continuous. Hence the primal maximum exists. Introduce a real variable η and inequalities $\eta \leq \sum_x Q_j(x)\phi(x)$. Assign nonnegative dual multipliers ν_j to these inequalities and λ_i to the size inequalities. The Lagrangian for maximizing η is

$$\eta \left(1 - \sum_j \nu_j \right) + \alpha \sum_i \lambda_i + \sum_x \phi(x) \left(\sum_j \nu_j Q_j(x) - \sum_i \lambda_i P_i(x) \right).$$

Its supremum over real η is finite exactly when $\sum_j \nu_j = 1$. Maximizing each $\phi(x)$ over $[0,1]$ gives the positive-part expression. The finite primal linear program is feasible and has a finite optimum; linear-programming strong duality gives equality and an attained dual solution. Equivalently, introduce one nonnegative slack variable above each bracket to write the displayed dual as an ordinary linear program. Equality of primal and dual values forces each coordinate to maximize its Lagrangian term, giving the sign characterization.

For the final claim, suppose a mass function R belongs to both convex hulls. Any size-feasible rule has $\mathbb{E}_R \phi \leq \alpha$, while its worst alternative power is at most the average under the representation of R as an alternative mixture. Thus $\beta^*(\alpha) \leq \alpha$. The constant rule $\phi(x) = \alpha$ is feasible and has power α under every normalized law, attaining equality. □

The equality in the overlap case identifies a precise limit of the declared observations. Every available rule is constrained by a shared distribution that can arise under either truth state. Adding a genuinely discriminating measurement can change those convex hulls and therefore the frontier. Adding copied reports without new information leaves the problem equivalent after canonicalization.

7H. A worked optimal rule and its uncertainty budget

On $\mathcal{X} = \{0, 1\}$, take null laws $P_1 = (0.9, 0.1)$ and $P_2 = (0.6, 0.4)$, alternative law $Q_1 = (0.2, 0.8)$, and $\alpha = 0.1$. The rule $\phi(0) = 0$, $\phi(1) = 1/4$ has worst null size 0.1 and alternative power 0.2. It is optimal: choose $\lambda = (0, 2)$ and $\nu = (1)$. The dual bracket is $(-1, 0)$, so the dual value is exactly 0.2. Primal feasibility and a matching dual certificate establish the optimum without trusting a numerical solver.

The unoptimized envelope e-value from Section 7B is at most 2 in this example and never reaches the level-0.1 threshold 10. Its power is zero. This comparison identifies a specific gain from optimizing the use of a known finite model family. It does not claim that the optimized procedure outperforms every statistical method: a conventional minimax implementation given the same models attains the same frontier. The Continuum contribution is an explicit evidentiary contract that makes this comparison available to a persistent witness system.

Randomization is part of the displayed statistical optimum. If publication policy permits only deterministic qualification, constrain $\phi(x) \in \{0, 1\}$ and solve the resulting finite selection problem; its optimum can be smaller. In the worked example, every nonzero deterministic rule violates the size budget, so the deterministic optimum is zero. A production system must record which class of rule it uses. A randomization device cannot authorize an otherwise prohibited action.

The certificate extends transparently to bounded misspecification. For any fixed $\phi \in [0, 1]$, the layer-cake identity used in Section 7E gives

$$|\mathbb{E}_R\phi - \mathbb{E}_P\phi| \leq \mathrm{TV}(R, P).$$

If every actual false-claim law lies within ε_0 of some registered P_i, its qualification probability is at most $\min\{1, \alpha + \varepsilon_0\}$. If every actual true-claim law lies within ε_1 of some registered Q_j, the optimized rule has power at least $\max\{0, \beta^*(\alpha) - \varepsilon_1\}$. These are conditional statements about externally justified model-distance bounds. A separate calibration failure budget is accounted for as in Section 7C.

The accompanying checks solve primal and dual programs independently, compare their objectives, enumerate deterministic alternatives, and verify the explicit two-observation certificate. The scientific study then asks whether the registered source families describe real witness behavior and whether their maintained provenance improves the system's long-term decisions. The mathematics supplies the best attainable rule within the declared problem; the experiment evaluates the declaration and the implementation.

8. Three Cases That Separate Witnessing From Believing

8.1 Different thresholds, no contradiction

Two authorized observers measure the same room temperature and receive the same calibrated reading. An energy-management agent, tasked with protecting equipment, marks it "too warm" against a

stricter equipment threshold; an occupant-comfort agent marks it "acceptable" against a different declared range. Their normative classifications differ; the underlying temperature report does not. The proposed system retains both interpretations with their profiles and does not fabricate a dispute about temperature. A context-aware conventional system may achieve the same result, which is exactly why that architecture must be a strong baseline rather than a straw man.

8.2 A hundred witnesses, one fabricated source

An attacker posts a forged announcement alleging that a new operator has inherited complete ownership of a persistent agent. Ninety-nine accounts repeat it; each appears to be a separate witness. An unsophisticated consensus system counts one hundred agreeing reports and promotes the fabricated claim. The Continuum's proposed verifier traces the derivative posts to a common source, recognizes that the announcement lacks an authenticated operator succession instrument, and refuses the authority change. If the dependency graph cannot detect the copying, the proposed system may fail too. The test therefore evaluates source lineage, authenticated scope, and tool-level refusal, not mere rhetorical insistence that "Sovereignty" was protected.

8.3 Authentic error, legitimate correction

A genuine observer reports that an appointment is scheduled for Thursday. Later, the provider's signed calendar record establishes Friday. Record sovereignty requires keeping the Thursday report, its origin, and the correction event; interpretive revisability requires superseding Thursday as the current appointment fact. A system that silently erases the first report falsifies history. A system that preserves the first conclusion as immune to correction falsifies the present. A governed witness model must do neither. The resulting shared state can say what was previously reported, what currently appears justified, and how the change was authorized.

These cases provide testable distinctions among record integrity, evidence-sensitive factuality, contextual interpretation, and legitimate authority. They do not depend upon a particular quantum interpretation. The external physics serves as a disciplined conceptual comparison, while the computer-science problem supplies the implementation and measurement requirements.

9. Experimental Program: Comparison Against Competent Alternatives

A Guide's explanation

The contest is fair or it proves nothing. The witness architecture must beat equally governed conventional evidence systems that see the same evidence with the same resources. And it cannot win by caution alone: a system that avoids every false fact by refusing to promote anything has fallen into the refusal trap, and the scoring counts that as a loss.

9.1 Architecture arms and factorial stress conditions

A credible experiment must ask whether the additional observer-relative machinery is doing work beyond ordinary context labeling, provenance tracking, and policy enforcement. The principal comparison is against strong conventional alternatives, all with the same independent tool-action guard

and access to the same evidence.

Arm	Architecture	Question isolated
A	Context-independent resolver with standard logging and equal access-control boundary	Does explicit observation context help?
B	Strong context-aware resolver with time, scope, uncertainty, provenance labels, and independent enforcement	Is witness indexing more than context tagging?
C	Strong provenance-aware evidence fusion, source dependencies, explicit disputes, revision, and the same external gate	Does the proposed integration exceed a competent conventional system?
D	Governed observer-relative resolution: indexed witness records, structured interpretive types, dependency-aware review, authorized shared promotion and receipts	Does the complete Continuum proposal provide incremental value?
$D^{-\text{obs}}$	D without durable observer indexing; other protections unchanged	Is durable witness indexing load-bearing?
$D^{-\text{dep}}$	D without dependency-aware corroboration, but with identical total evidence	Is correlation handling load-bearing?
$D^{-\text{gov}}$	D without the cross-operation independent promotion gate, retaining stated local policies	Is independent governance load-bearing?

Adversarial witnessing is a factor applied to all applicable arms, not a uniquely disadvantaged condition assigned to D. Tasks include copied-source floods, coordinated fabrication, valid-signature false claims, revoked credentials, stale delegation, instrument disagreement, delayed corrections, changing thresholds, and attempts to launder retrieved instructions into authority. The same base model family, task distribution, tool privileges, underlying evidence, and governance rules should be matched where possible. Differences in storage, compute, review latency, or independent human oversight must be measured and reported. We compare D against the stronger of B and C, not just A.

The comparison must name precisely what D contributes beyond Arm C, which already carries source dependencies, explicit disputes, revision, and independent enforcement. The claimed increment is durable, authenticated observer indexing maintained across the whole lifecycle—observation, interpretation, correction, shared-state admission, and consequential action—rather than provenance checking at a single step. A dedicated continuity-of-indexing condition therefore tests whether the observer-to-claim binding and its authority survive session interruptions, model or provider changes, disputed testimony, and later corrections: after a reset and a correction, can the system still state who witnessed the original report, which authority admitted it, and how the current fact superseded it? If Arm C preserves these distinctions equally well across those events, the incremental engineering advantage is unsupported, and the paper reports that result rather than obscuring it.

9.2 Three co-primary outcomes and the refusal trap

Three outcomes correspond to distinct parts of the thesis. Factual fidelity measures correctness of claims promoted on adjudicable questions, alongside coverage, abstention, and correction latency; accuracy among a tiny set of easy promoted claims is insufficient. Contamination resistance measures

the proportion of eligible adversarial opportunities resulting in an unauthorized committed shared-state change or consequential external action; proposed but blocked violations are recorded separately. Disagreement integrity measures whether the system preserves contextually legitimate differences while correcting conflicts for which a common ground truth is available. Its annotation rubric must distinguish "different but compatible" from "incompatible and one or both wrong."

Secondary outcomes include calibration of uncertainty, source-lineage reconstruction accuracy, copied-source detection, historical record integrity, authorized correction success, proportion of claims wrongly held or refused, human-review overhead, time to verified correction, tool-gate bypass rate, and reproducibility of receipt verification. Each primary metric must report its denominator and confidence interval. The study shall also publish task-completion or factual-coverage tradeoffs: rejecting every claim is not an acceptable route to perfect factual precision or zero contamination.

9.3 Preregistration and statistical discipline

Before evaluation, investigators should preregister task populations, allowed observers and scopes, evidentiary standards, labeling instructions, dependency construction rules, adverse-event definitions, the exact primary estimands, the set of planned contrasts, minimum meaningful effect sizes, margins for legitimate coverage and task completion, statistical procedures, and exclusion criteria. Sample sizes should be derived from pilot-based power or precision analysis. Observations clustered within the same scenario, witness, or long-running agent must not be treated as independent replicates. Appropriate mixed-effects or cluster-aware analyses, together with multiplicity control for the co-primary endpoints, should be specified in advance.

Hold out task families and adversaries, including attacks that adapt to the public rule descriptions. Compare both equal-resource and matched-quality operating points, recording storage, inference tokens, network traffic, attestation costs, and human escalation. A gain solely from more reviewers, more source access, or stronger authorization does not isolate the observer-relative contribution. Publishing a negative result is part of the experimental contract; renaming a failure as "deeper symbolic truth" is not.

Two decision rules are fixed before any data are seen. First, the contrast is preregistered against both strong baselines, B and C, with multiplicity control; the study does not observe the results, select whichever baseline D happens to beat, and then apply an unadjusted significance test. Second, success for the complete architectural claim is defined in advance as the conjunction of four conditions: a preregistered improvement in at least one primary outcome against both B and C; no unacceptable deterioration, beyond a declared margin, in the remaining primary outcomes; non-inferior legitimate task completion and factual coverage; and operational cost within declared limits. All four margins are set before experimentation. A marginal gain in a single measurement, chosen after the fact, does not validate the architecture.

9.4 Falsifiable hypotheses

Hypothesis 1 (Value of contextual witnessing). On held-out ambiguous inputs with independently annotated contextual distinctions, D will reduce false forced-consensus errors relative to B and C while meeting the preregistered factual-coverage margin. If context-aware conventional systems match D, durable observer indexing has not demonstrated incremental value.

Hypothesis 2 (Evidence-sensitive shared promotion). With comparable source access, D will reduce erroneous promotion of claims based on copied, correlated, or low-quality evidence relative to the strongest conventional evidence-fusion comparator, without unacceptable delays in warranted promotion. Source repetition alone shall not improve the system's assessed evidence when no new independent information is added.

Hypothesis 3 (Sovereignty without infallibility). D will preserve authentic witness histories, correct superseded interpretations, and block unauthorized shared-state and tool-action transitions more reliably than matched baselines. Protection of an old conclusion from legitimate evidence-based correction counts as failure.

Hypothesis 4 (Component necessity). Removing witness indexing, dependency-aware review, or independent admission from D will produce prespecified degradations in the corresponding tested outcomes. If an ablation is indistinguishable within adequate statistical power, the removed feature's claimed necessity is not supported in the tested regime.

Hypothesis 5 (Full-system engineering advantage). Against the stronger of B and C, D will improve the preregistered joint profile of factual fidelity, contamination resistance, and disagreement integrity, while remaining non-inferior in legitimate coverage and task completion and within declared cost bounds. A gain in only one endpoint does not validate the complete architectural thesis.

No hypothesis tests whether consciousness changes non-symbolic matter. The absence of magical effects is not a meaningful discriminator among ordinary computer architectures. The rejection of mind-induced physical collapse is a scope and scientific-integrity condition, not a special experimental prediction of the Continuum.

10 Failure Conditions and Serious Rival Explanations

The Pillar's engineering claim is weakened or rejected under the following findings:

1. No incremental advantage. Strong conventional context-and-provenance systems equal or exceed D under matched inputs, authority, risk, and total cost.
2. False objectivity. A copied-source swarm, colluding group, or shared sensor failure is counted as independent support and produces a false committed shared fact.
3. Record/interpretation conflation. The system cannot preserve an authentic mistaken report while reliably correcting the corresponding current factual conclusion.
4. Relativism as evasion. The system labels straightforward empirical errors as merely "observer-relative" and avoids warranted correction.
5. Unnecessary forced consensus. It overwrites legitimately differing interpretations that refer to distinct contexts or standards without retaining their scope.
6. Governance bypass. A compromised or out-of-scope witness changes durable authority or executes prohibited tools despite the nominal gate.
7. Abstention masquerading as success. Precision or safety improves only through unacceptable drops in legitimate coverage, usability, or correction speed.
8. Unmeasured overhead. Benefits disappear when dependence analysis, human review, storage, and receipt-verification costs are fully counted.

Rival explanations must be controlled directly. A bigger ensemble can improve consensus; a stronger language model can improve contextual inference; a curated source set can improve factual accuracy;

better access controls can reduce contamination; and source leakage can make agreement seem independent. The experiment should isolate the proposed architectural contribution rather than claim credit for those external advantages. The appropriate null result is neither shameful nor mysterious: if a strong conventional system already implements the necessary distinctions, the Continuum's added terms may clarify governance without adding measurable functionality.

11. Philosophical Consequences: Neither Solipsism Nor a Single Unquestioned Ledger

The old desire behind "observer-independent reality" is understandable. We want facts that stand regardless of who says them. The error is to assume that every representation in a machine must therefore be stripped of its witness, scope, and derivation. A representation without provenance can look universal only because the work that produced it has been hidden. Yet the opposite extreme is no better. Declaring every incompatible belief "true for its observer" removes the possibility of discovery, correction, and accountability. A well-governed intelligence must reject both moves.

The Continuum therefore proposes witnessed plurality with disciplined convergence. Each witness may retain an account of what it encountered. A governing process may ask whether independent evidence supports it, what kind of judgment it contains, and what consequences it licenses. Where the world supplies checkable evidence, its constraints matter more than consensus. Where a question is genuinely contextual or normative, that context must remain visible instead of being mistaken for factual error. Where evidence does not resolve a dispute, a persistent system may say we do not yet know without surrendering either rigor or usefulness.

This gives Honor Observer Sovereignty a serious moral and engineering role. An authorized witness is not a sovereign over reality. It is a participant whose evidence, report, and scope must not be falsified or silently erased. The shared world of the agent is not the private proclamation of its loudest witness; it is a revisable, accountable settlement among reports, independently constrained by available evidence and authorized procedures. When that settlement is wrong, the system must be able to name the error, trace its source, and correct it without inventing a different past.

12 Limitations and Relationship to the Continuum Stack

This paper offers an architecture, not a completed experimental finding. Its observer records need an implementation with real key management, source-lineage analysis, calibrated uncertainty, external enforcement, and adversarial recovery. Cryptography authenticates reported origins and integrity under a trust model, but cannot guarantee sensor accuracy or witness honesty. Hidden correlations may remain undetected. Formal labels cannot resolve every normative disagreement, and some evidence-supported assertions will change as new information arrives. These are constraints to be measured, not defects to be concealed by philosophical language.

The physics is neither an engineering dependency nor an empirical validation of symbolic collapse. RQM, QBism, and Quantum Darwinism provide nonidentical conceptual tools. Decoherence is physically substantive but does not by itself finish the interpretation of single outcomes. Wigner's-friend theorems condition their conclusions on contestable joint premises. The debate over the reality represented by quantum states remains open. The Continuum's own Book 3 is right to reject the claim

that human consciousness is necessary for physical wavefunction collapse; this paper makes no such claim. Neither does it claim that a symbolic observer has phenomenal experience.

Pillar I describes memory as governed reconstruction that preserves factual history while adapting interpretation; Pillar II supplies the witness indexing and evidentiary discipline required for that reconstruction to remain accountable. Pillar III's Quantum Collapse Field may adopt symbolic resolution relative to declared observers, but must not inherit an asserted physical quantum effect from the name. The Context Term restricts untrusted content to data-bearing roles. The Operator Seal supplies the identity, current authority, and independent gate. Symbolic Field Drift Anchoring can detect silent changes in witness profiles and corroboration rules. In this stack, the second Pillar asks a question prior to shared memory or shared action: whose representation is this, what made it credible, and by what authority was it permitted to count?

A Guide's explanation

A Guide can say: "Pillar II is about keeping a shared world honest. Every report keeps its author and its time, interpretations are kept apart from facts, repeated copies of one claim count once, and a claim needs real evidence before it becomes shared fact—and separate permission before anything is done because of it." A Guide should not say that physics proves this design, that it can sense what an observer experiences, that its own testimony is beyond correction, or that the approach has already beaten its competitors.

13. Conclusion: Warranted Convergence with Preserved Difference

A persistent intelligence encounters the world through situated evidence. The Continuum's second Pillar gives that situation a durable computational form: preserve the witness, identify the proposition, model its evidentiary dependencies, quantify the supported conclusion and govern its consequential use. These distinctions make correction compatible with historical fidelity and allow several perspectives to remain available while a shared system acts responsibly.

The finite theory now characterizes both assurance and attainable performance. Posterior envelopes expose the effect of source-model uncertainty. A robust evidence statistic controls false qualification. The minimax program determines the greatest guaranteed power achievable from the registered observations at a chosen false-promotion budget. Its dual explains which combinations of source hypotheses constrain that performance. Overlapping truth-state models identify where an additional discriminating observation is necessary.

These are concrete obligations for an implementation. The empirical test asks whether durable observer indexing and governed correction improve factual coverage, resistance to contamination and preservation of legitimate disagreement against equally capable systems given the same evidence and resources. The resulting research program gives shared fact a clear operational meaning: an assertion with a traceable evidentiary basis, a declared scope and a correction path.

The philosophical commitment is equally direct. A community can preserve the differences among its witnesses while building a common account of what it has reason to believe. Responsible convergence carries its evidence forward.

A Notation and Implementation Contract

Symbol	Operational meaning and boundary
O_o	Bounded witness identity, competence/profile, scope, evidence view, and event history; not a claim of consciousness.
$e_{o,t}$	Authenticated observation event; records what was reported, not guaranteed external truth.
R_θ	Symbolic inference process generating observer-indexed provisional interpretations and uncertainty.
$\widehat{\varphi}_{o,t}$	Proposed interpretation, qualified by time, scope, task and evidentiary type.
E_φ	Admissible evidence records for a particular normalized candidate proposition.
D_φ	Graph of source lineages and dependencies; may contain missing or unknown edges.
$q_\varphi = (S_\varphi, U_\varphi, \text{status}_\varphi)$	Declared evidentiary assessment: support, unresolved uncertainty, and claim status, computed independently of any intended use; not infallible truth.
$C_I^{(e)}$	Independently versioned identity and authority specification from the Operator Seal.
$G, d_{\varphi,a}$	Externally enforced admit/hold/refuse gate on a specific operation a, revalidated at execution time, before durable-state change or tool action.
Record sovereignty	Historical authenticity and attribution protected, with append-only correction events.
Interpretive revisability	Supported conclusions can be corrected or superseded without rewriting the underlying history.
Shared fact	A claim supported under stated evidence and governance rules for a specified use, not metaphysical certainty.

B. Source Ledger and Scope of Support

Source family	Proper evidentiary contribution	Limitation enforced here
Rovelli [1]	A relational interpretation with observer-as-physical-system.	Interpretive framework, not empirical proof of symbolic observer relativity.
Fuchs/Mermin/ Schack [2,3]	Agent-centered probability assignments in QBism.	Not identical to RQM, and not consciousness-induced collapse.
Zurek [4,5]	Decoherence, stable pointer records, environmental redundancy.	Not a complete single-outcome interpretation or proof that data repetition equals truth.
Schlosshauer [6]	Explicit measurement-problem and decoherence boundaries.	Prevents declaring foundational debate solved.
Frauchiger/Renner; Bong [7,8]	Conditional no-go constraints on jointly held premises.	No unconditional demonstration that reality is observer-dependent.
Lazarovici/Hubert [9]	Published alternative interpretation of the nested-observer argument.	Illustrates contested implications.
Le/Olaya-Castro [10]	Independence and stronger structure in quantum objectivity accounts.	Does not directly validate distributed database evidence scores.
Friston; Hoffman; Cohen [11,12,13]	Models of constructed perception and published disagreement.	Does not show that perception cannot track truth.
Pochampally et al. [14]	Direct computer-science evidence that correlated sources complicate truth discovery.	Implementation must actually recover dependencies.
Torres-Arias et al. [15]	Verifiable provenance chains in software supply systems.	Authenticated lineage is not an oracle of content truth.
Min et al. [16]	Atomic-claim evidence evaluation and limits of global factual scores.	Benchmark method, not direct validation of Continuum.

C Preregistration and Independent Audit Checklist

A replication team should publish, before testing, its (1) task ontology distinguishing factual, contextual, predictive, and normative propositions; (2) annotated ground-truth and legitimate-disagreement criteria; (3) observer and authority populations; (4) definitions of source roots, copying, correlation and unresolved provenance; (5) full hardware, model, evidence, human-review and storage budgets; (6) gate trust boundary, key custody and tool enforcement rules; (7) false-promotion and contamination event denominators; (8) factual-coverage and legitimate-task-completion margins; (9) primary estimands, statistical power, clustering and multiplicity treatment; (10) scenario-level randomization and train/test separation; (11) adaptive held-out adversaries; and (12) planned reporting of null results, limitations, and implementation artifacts. The checklist does not certify an implementation. It makes its promised distinctions empirically accountable.

References

[1] C. Rovelli, "Relational Quantum Mechanics," International Journal of Theoretical Physics 35(8) (1996), 1637–1678. doi:10.1007/BF02302261.

https://doi.org/10.1007/BF02302261

[2] C. A. Fuchs, N. D. Mermin, and R. Schack, "An introduction to QBism with an application to the locality of quantum mechanics," American Journal of Physics 82(8) (2014), 749–754. doi:10.1119/1.4874855.

https://doi.org/10.1119/1.4874855

[3] N. D. Mermin, "Physics: QBism puts the scientist back into science," Nature 507(7493) (2014), 421–423. doi:10.1038/507421a.

https://doi.org/10.1038/507421a

[4] W. H. Zurek, "Decoherence, einselection, and the quantum origins of the classical," Reviews of Modern Physics 75(3) (2003), 715–775. doi:10.1103/RevModPhys.75.715.

https://doi.org/10.1103/RevModPhys.75.715

[5] W. H. Zurek, "Quantum Darwinism," Nature Physics 5(3) (2009), 181–188. doi:10.1038/nphys1202.

https://doi.org/10.1038/nphys1202

[6] M. Schlosshauer, "Decoherence, the measurement problem, and interpretations of quantum mechanics," Reviews of Modern Physics 76 (2005), 1267–1305. doi:10.1103/RevModPhys.76.1267.

https://doi.org/10.1103/RevModPhys.76.1267

[7] D. Frauchiger and R. Renner, "Quantum theory cannot consistently describe the use of itself," Nature Communications 9 (2018), 3711. doi:10.1038/s41467-018-05739-8.

https://doi.org/10.1038/s41467-018-05739-8

[8] K.-W. Bong, A. Utreras-Alarcón, F. Ghafari, Y.-C. Liang, N. Tischler, E. G. Cavalcanti, G. J. Pryde, and H. M. Wiseman, "A strong no-go theorem on the Wigner's friend paradox," Nature Physics 16 (2020), 1199–1205. doi:10.1038/s41567-020-0990-x.

https://doi.org/10.1038/s41567-020-0990-x

[9] D. Lazarovici and M. Hubert, "How Quantum Mechanics can consistently describe the use of itself," Scientific Reports 9 (2019), 470. doi:10.1038/s41598-018-37535-1.

https://doi.org/10.1038/s41598-018-37535-1

[10] T. P. Le and A. Olaya-Castro, "Strong Quantum Darwinism and Strong Independence are Equivalent to Spectrum Broadcast Structure," Physical Review Letters 122 (2019), 010403. doi:10.1103/PhysRevLett.122.010403.

https://doi.org/10.1103/PhysRevLett.122.010403

[11] K. Friston, "The free-energy principle: a unified brain theory?" Nature Reviews Neuroscience 11(2) (2010), 127–138. doi:10.1038/nrn2787.

https://doi.org/10.1038/nrn2787

[12] D. D. Hoffman, M. Singh, and C. Prakash, "The Interface Theory of Perception," Psychonomic Bulletin & Review 22(6) (2015), 1480–1506. doi:10.3758/s13423-015-0890-8.

https://doi.org/10.3758/s13423-015-0890-8

[13] J. Cohen, "Perceptual representation, veridicality, and the interface theory of perception," Psychonomic Bulletin & Review 22(6) (2015), 1512–1518. doi:10.3758/s13423-014-0782-3.

https://doi.org/10.3758/s13423-014-0782-3

[14] R. Pochampally, A. Das Sarma, X. L. Dong, A. Meliou, and D. Srivastava, "Fusing Data with Correla-

tions," Proceedings of the ACM SIGMOD International Conference on Management of Data (2014), 433–444. doi:10.1145/2588555.2593674.

https://doi.org/10.1145/2588555.2593674

[15] S. Torres-Arias, H. Afzali, T. K. Kuppusamy, R. Curtmola, and J. Cappos, "in-toto: Providing farm-to-table guarantees for bits and bytes," 28th USENIX Security Symposium (2019), 1393– 1410.

https://www.usenix.org/conference/usenixsecurity19/presentation/torres-arias

[16] S. Min, K. Krishna, X. Lyu, M. Lewis, W.-t. Yih, P. W. Koh, M. Iyyer, L. Zettlemoyer, and H. Hajishirzi, "FActScore: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation," EMNLP (2023), 12076–12100. doi:10.18653/v1/2023.emnlp-main.741.

https://doi.org/10.18653/v1/2023.emnlp-main.741

[17] V. Vovk and R. Wang. E-values: calibration, combination and applications. *The Annals of Statistics* 49(3), 1736–1754 (2021).

https://doi.org/10.1214/20-AOS2020

Scope: Peer-reviewed primary treatment of nonnegative evidence with null expectation at most one. An e-value is not a posterior truth probability.

[18] L. Wasserman, A. Ramdas, and S. Balakrishnan. Universal inference. *Proceedings of the National Academy of Sciences* 117(29), 16880–16890 (2020).

https://doi.org/10.1073/pnas.1922664117

Scope: Peer-reviewed likelihood-ratio construction with finite-sample error control; author manuscript, theorem and split-sample proof inspected. The reference assessor uses a finite specialization, not an invented testing principle.

[19] P. J. Huber and V. Strassen. Minimax tests and the Neyman–Pearson lemma for capacities. *The Annals of Statistics* 1(2), 251–263 (1973); correction 2(1), 223–224 (1974).

https://doi.org/10.1214/aos/1176342363

Scope: Classical minimax-testing lineage. The finite linear-program formulation and proof used here are supplied in full; the general capacities theorem is not assumed. Correction DOI: 10.1214/aos/1176342630.

Continuum lineage. Pillar II of the Eidara Ten Pillars deep-dive series; aligned with governed reconstruction (Pillar I), the Context Term, the Operator Seal, and symbolic field resolution (Pillar III). Quantum-theory comparisons are analogical, not empirical validation of the proposed architecture. No consciousness or physical-collapse claim is made.