

Update Integrity Standard (UIS)

A structural ethic for preserving corrigibility in record-based systems

Jeremy C. Jones (ORCID 0009-0007-2515-3774)

HoldingLight LLC

Version v1.0 • 2026-04 • CC BY 4.0

universalcollapse.com

© 2026 Jeremy C. Jones — HoldingLight LLC

This document is a versioned internal standard of the Universal Collapse Theory research program, published for transparency, reuse, and critique. It is not presented as a field-ratified standard.

Abstract

Any system that learns from its own outcomes—scientific communities, institutions, minds, algorithms—depends on an update loop: outcomes produce records, and records modify future constraints. This standard identifies the structural conditions under which that loop remains corrigible (update integrity) and names the common failure modes that break it: record falsification, pseudo-redundancy, constraint freezing, selective update rules, coercive agreement, discriminator-free coherence, and identity binding. For each failure mode, a corresponding repair protocol and diagnostic indicator are specified. This version specifies measurement proxies for corrigibility, signal-state protocols for neutrality and commitment (referencing the S_1 – S_3 stabilization family), an Empirical Ledger template for per-domain tracking, and a Minimal Implementation appendix specifying measurement thresholds. UIS is a companion to *The Structuralization of Empiricism* (Jones, 2026b), which introduces the recursive kernel and stabilization signatures that UIS operationalizes, and to *Records Across Nature, Life, and Mind* (Jones, 2026a), which defines the persistence layer UIS protects.

Keywords: update integrity; corrigibility; record discipline; redundancy audit; discriminators; reporting standards; research integrity; philosophy of science

Purpose

The Update Integrity Standard (UIS) formalizes a single ethic implied by a record-based collapse framework: do not corrupt the update loop. UIS is not a social virtue code (“be nice”). It is a structural requirement for any system that aims to remain coherent over time—scientific communities, organizations, minds, and institutions alike.

Core principle

A system remains update-capable when (i) records are not falsified, (ii) redundancy is not counterfeited, and (iii) constraints remain revisable under legitimate evidence. “Legitimate evidence” here means evidence admitted under pre-declared standards, with inspectable provenance, symmetric evaluation rules, and enough independence information to assess redundancy quality.

Review target

UIS does not ask the reader to accept Universal Collapse Theory as a metaphysical system or to treat this document as a field-ratified compliance standard. It asks whether update integrity can be operationalized as a practical standard for record-bearing systems: preserve records, audit redundancy, specify discriminators, keep constraints revisable, and report the current stabilization mode. UIS should be accepted provisionally only if its corruptor classes, indicators, repair protocols, reporting standard, and empirical ledger improve the detection and repair of update-loop failure. It should be revised or rejected if these procedures do not distinguish corrigible update from pseudo-redundancy, constraint freezing, selective update, or locked stabilization better than ordinary domain-local practice.

Triad placement

This document occupies the update-integrity position in the UCT standards layer. *Records Across Nature, Life, and Mind* defines the persistence layer: what records are and why they make collapse cumulative. *The Structuralization of Empiricism* defines the stabilization layer: how record-bearing empirical systems converge, delay, or lock under constraints. UIS defines the governance layer: how update loops remain corrigible through record discipline, independence audit, symmetric standards, pre-declared discriminators, and explicit rollback/correction procedures.

UCT kernel context (informal)

$$(\Omega, K) \rightarrow C^K(\Omega, K) = x^* \rightarrow (R, S, T) \rightarrow K' = U(K, x^*, R)$$

Here S denotes residue or loss/side-effect structure and T denotes record-time or ledger index; UIS focuses on R, U, and redundancy quality because those are the interfaces most directly involved in corrigibility.

UIS protects three interfaces in this loop:

- Records R (what happened and what is recoverable)
- Update map U (how records are allowed to change constraints)
- Redundancy quality (whether multiple records are genuinely independent)

Definitions

Update loop

The update loop is any process by which outcomes and their records modify future constraints—belief revision, learning rules, institutional reform, scientific error correction, or model retraining.

Update integrity

Update integrity means the loop remains corrigible: false models can be corrected by reality-contact and by audit of records. Integrity is graded, not binary; UIS measures and improves it.

Constraint set (K)

The active rules, assumptions, thresholds, policies, priors, or operating conditions that shape what outcomes are allowed, selected, interpreted, or acted on. K is what records modify through U.

Effective independence (k_{eff})

The estimated number of genuinely independent records after shared sources, shared instruments, shared incentives, copied evidence, or correlated errors are discounted. Reported in independence audits and UIS-RS disclosures.

Record (R)

A durable, inspectable trace of an outcome that can constrain future updates. Examples include datasets, instrument logs, publications, audit trails, and memory traces. The full cross-domain definition is given in *Records Across Nature, Life, and Mind* (Jones, 2026a).

Discriminator

A pre-specified observation, failure condition, or comparison that would distinguish a claim from a plausible alternative and force update, scope reduction, or withdrawal. Discriminators are required for Level 3 claims under UIS-RS.

Corruptors

UIS names the common failure modes that break corrigibility. Each corruptor targets R, U, or redundancy.

- **Record falsification:** Lying, data fabrication, selective deletion, retroactive rewriting of what was predicted or observed.
- **Pseudo-redundancy:** Manufacturing the appearance of consensus via correlated sources, echo chambers, pressure, or copy-pasted evidence shards.
- **Constraint freezing:** Dogma: treating provisional constraints as non-negotiable axioms; refusing to define what would change the belief/policy.
- **Selective update rules:** Asymmetric standards: evidence counts only when it supports the preferred conclusion; logic is applied unevenly.
- **Coercive agreement:** Forcing public agreement without independent access to records; compliance replaces consensus. In UCT terms, this is coercive collapse.
- **Discriminator-free coherence:** Claims insulated from discriminators; emotive or aesthetic coherence replaces constraint coherence. Informally: vibe-metaphysics.
- **Identity binding:** Beliefs become self-binding; updating feels like self-destruction, so U is effectively disabled in that channel. Distinct from social-psychological identity fusion (Swann et al.,

2012), where group membership becomes personally fused with the self; UIS uses “identity binding” in an update-loop sense.

Integrity indicators

Use these diagnostics to audit whether a loop is generative (corrigible) or supplanted (self-sealing):

- **Pre-commitment:** the system can state in advance what observations would change its conclusion.
- **Audit trail:** predictions, decisions, and justifications are recorded before outcomes are known.
- **Independent replication:** redundancy comes from genuinely independent records (or an explicit k_{eff} estimate). This usage aligns with the reproducibility/replicability distinction in the National Academies of Sciences, Engineering, and Medicine report (NASEM, 2019): reproducibility under the same data and analytic conditions is not the same integrity signal as independent studies obtaining consistent results with their own data.
- **Error admission is low-cost:** acknowledging error does not trigger identity or status collapse.
- **Symmetric standards:** the same evidentiary and logical filters apply regardless of conclusion.
- **Ablations exist where causal learning/adaptation is claimed:** disabling the proposed update mechanism should reduce the claimed performance gain or stabilization effect (causal attribution).

Measurement proxies for corrigibility

Corrigibility is graded. The indicators above describe qualitative conditions; the following proxies quantify them where measurement is feasible. Each proxy operationalizes update sensitivity: the degree to which records are permitted to modify K .

- **Replication success rate:** Proportion of verified-independent attempts that replicate a claimed finding under preserved independence conditions. Low rates indicate weak S_1 or corrupted R .
- **Time-to-retraction:** Elapsed time from confirmed disconfirmation to formal correction of the record. Long latencies indicate weak U . In publication contexts, this proxy should be interpreted in the corrective sense used by the Committee on Publication Ethics (COPE, 2025): retraction protects the integrity of the record rather than serving primarily as punishment.
- **Correction lag:** Time from first credible disconfirming evidence to effective constraint update (retraction, model revision, policy change). Distinguishable from time-to-retraction by tracking the *evidence-to-update* interval rather than disconfirmation-to-formality.
- **Variance reduction across independent trials:** Rate at which cross-trial variance decays as independent records accumulate. Failure to reduce variance under preserved independence indicates pseudo-redundancy or update insensitivity.
- **Rollback completion rate:** Proportion of failed, retracted, or integrity-compromised claims that are visibly corrected, scoped down, withdrawn, or replaced within a pre-declared window. Useful where formal retraction is not the native unit of correction (institutions, algorithmic systems, policy regimes).

These proxies are not exhaustive. They are default starting points for domains where no native corrigibility metric exists. Domain-specific metrics may substitute or extend them; substitutions must preserve the property measured (update sensitivity under disconfirmation).

Repair protocols

UIS is practical: it specifies small, repeatable interventions that restore corrigibility without requiring total worldview replacement.

- **R-Ledger (Record discipline):** Write predictions and rationales with timestamps; preserve failures; forbid retroactive edits (append only). Where records are digital datasets or computational artifacts, R-Ledger discipline should be read alongside FAIR data principles (Wilkinson et al., 2016): records must be findable, accessible, interoperable, and reusable enough to support later audit and update.
- **Discriminator rule:** For every nontrivial claim, name at least one observation that would discriminate it from a plausible alternative.
- **Independence audit:** Estimate whether sources are independent; downgrade confidence when redundancy is correlated (report k_{eff}).
- **Reversibility first:** Prefer reversible, low-stakes tests before irreversible commitments (prevents identity binding).
- **Attack-to-Stack policy:** Replace adversarial language with stack placement: argue level and interface before arguing content. (“Stack” here means the level/interface at which a claim operates: record, constraint, update rule, signal state, or domain application.)
- **Constraint loosening:** If a belief/policy is frozen, reduce grip by one notch: hold it provisionally and run a bounded test.
- **Post-mortem update:** After outcomes resolve, log what constraint was missing and what update will be made to K (not just what happened).

Neutrality and commitment protocols

Every update loop reaches points where it must decide: remain neutral, or commit to a constraint refinement. The Structuralization of Empiricism (Jones, 2026b) shows that both moves are structurally grounded in the signal state of the system, not merely in epistemic virtue or personal disposition. UIS operationalizes these grounds as decision protocols.

Neutrality and commitment errors are symmetric. A system that remains neutral when S_1 has converged under verified independence is as malformed as one that commits when S_2 has not plateaued. Each direction corrupts the update loop in a characteristic way: premature commitment manufactures false S_1 and locks K against legitimate disconfirmation; delayed commitment freezes the loop in symmetric indecision and wastes update capacity.

Signal-state preconditions

Neutrality is structurally warranted when:

- Redundancy is low or unaudited (independence not yet verified).
- Competing hypotheses remain symmetric under current K (S_2 active).
- Sweep tests have not yet been applied, so S_3 status is unknown.
- Prior-constraint strength is insufficient to resolve selection C^K without structural distortion.

Commitment is structurally warranted when:

- Redundant independent records have converged (S_1 has resolved under verified independence).
- Neutrality-induced delay no longer alters outcome (S_2 plateau).
- Sweep tests reveal bounded, acceptable hysteresis (S_3 controlled).
- Update integrity conditions remain satisfied (R durable, U sensitive, standards symmetric).

Stopping criteria

Stopping rules must be pre-declared, not retrofit. Acceptable stopping triggers include: a pre-registered threshold for S_1 convergence; a fixed budget of replications, trials, or iterations; an S_2 plateau confirmed across at least two independent resolution attempts; a bounded S_3 loop area below a domain-specified limit. Stopping without pre-declared triggers is a corruptor (selective update): the system retroactively chooses the stopping point that supports the preferred conclusion. This is not merely procedural: undisclosed flexibility in data collection, analysis, and reporting can inflate false-positive rates (Simmons, Nelson, & Simonsohn, 2011), which is why preregistered stopping rules function as update-integrity safeguards. Where pre-declaration is infeasible because the domain is genuinely exploratory, UIS-RS must label the claim exploratory and prohibit Level 3 commitment until a confirmatory test is specified.

Multi-level alignment

Commitment does not require alignment at every level simultaneously. A claim may be committed at the architectural level (Level 1) while remaining neutral at the divergence level (Level 3). What must not occur is asymmetric collapse: committing at Level 3 (a specific empirical prediction) while leaving Level 1 scope undefined, or committing at Level 1 (a structural pattern) while claiming Level 3 predictive force without signal-state support. Each level's commitment has its own preconditions; mixing them corrupts the update loop.

Stabilization-mode classifications

Systems can be classified by which signal-state configuration they occupy:

- **Converging:** S_1 active and producing variance reduction; commitment conditions approaching.
- **Symmetric:** S_2 dominant; neutrality appropriate; stopping rules in force.
- **Path-dependent:** S_3 signatures present; requires independence audit before commitment.
- **Locked:** S_3 present without S_1 convergence and without update sensitivity; pathological stabilization; requires repair protocols.
- **Drifting:** No stable signal state; neither S_1 convergence nor S_2 plateau; may indicate missing K (constraint set incomplete) or failed R (records not durable).

Classification is not a verdict. It is a snapshot: the same system may move between modes as records accumulate, integrity is preserved or corrupted, or K is revised. Reporting the current mode under UIS-RS (below) disciplines claims about the system's epistemic state.

UIS Reporting Standard (UIS-RS)

Minimum disclosure when you claim a belief, model, or policy is "update-capable": UIS-RS is structurally adjacent to open-science reporting frameworks such as the Transparency and Openness

Promotion (TOP) Guidelines (Nosek et al., 2015), which treat transparency, data availability, preregistration, and reporting standards as ways to increase the verifiability of empirical claims.

- **Claim boundary:** what domain is this operating in: physical, biological, cognitive, institutional, or algorithmic? Is the claim about records, constraints, update rules, signal state, or domain application?
- **Record set:** what records exist, where they live, and what is preserved (R).
- **Redundancy:** how many independent records support the claim; what is the estimated k_{eff} ?
- **Discriminators:** what would change the conclusion; what alternatives were considered?
- **Update rule:** what is allowed to update in K, how fast, and under what triggers (U).
- **Integrity risks:** which corruptors are most likely here (e.g., pseudo-redundancy, identity binding) and what safeguards are in place.
- **Stabilization mode:** which mode applies (converging, symmetric, path-dependent, locked, drifting); what signal states justify the classification.
- **Claim level:** Level 1 architectural mapping, Level 2 interpretive lens, or Level 3 testable claim.
- **Rollback trigger:** what observation, replication result, audit failure, or integrity breach would force correction, scope reduction, or withdrawal?
- **Stopping rule:** what pre-declared threshold, replication budget, S_2 plateau, or S_3 bound justifies commitment?

Scope and application

UIS applies wherever a system's future behavior depends on records of its past outcomes. This includes empirical research communities (where update integrity maps onto replication standards, preregistration, and error correction), cognitive and perceptual systems (where beliefs and expectations update under incoming evidence), biological systems (where molecular records constrain developmental and adaptive trajectories), institutional and policy systems (where accumulated precedent shapes future decision rules), and algorithmic systems (where training data and feedback loops modify model constraints). The standard is domain-portable: it does not prescribe what a system should believe, only the structural conditions under which its update loop remains corrigible. For the formal epistemological framework that motivates these conditions, see *The Structuralization of Empiricism* (Jones, 2026b).

UIS does not certify that a belief, model, policy, or institution is correct. It provides a disclosure standard for whether the update loop has specified its records, redundancy structure, discriminators, update rule, integrity risks, and stabilization mode. Domain-specific standards remain authoritative for domain-specific truth, safety, legality, and ethics.

One-line carryforward

Don't corrupt the update loop: protect records, preserve independent redundancy, and keep constraints revisable.

Appendix A: The Empirical Ledger

To prevent scope creep, applications of the Structuralization of Empiricism framework (and of update-integrity standards generally) should be logged explicitly using the following template. Each domain application completes all fields.

Field	Entry
Domain	(e.g., pharmacology, climate science, machine learning)
Ω specification	What are the distinguishable alternatives?
K specification	What constraints are active? (instruments, priors, protocols)
Record type (R)	What traces persist? (datasets, logs, publications)
Independence audit method	How is independence verified? (different labs, different data sources, k_{eff} reported)
S₁ status	Does redundancy drive convergence? (measured / expected / not tested / failed)
S₂ status	Does neutrality delay resolution? (measured / expected / not tested / failed)
S₃ status	Does constraint sweep reveal hysteresis? (measured / expected / not tested / failed)
Control conditions satisfied	List controls applied and any uncontrolled confounds.
Update integrity assessment	Are records preserved? Is U sensitive to R? Are standards symmetric? Stabilization mode?
Correction/rollback history	Log any revisions, retractions, or scope changes since initial entry.
Claim level	Level 1 (architectural mapping), Level 2 (interpretive lens), or Level 3 (testable claim).
Discriminator / rollback trigger	What observation, failure, replication result, or integrity audit would change, constrain, retract, or roll back the claim?
Stopping rule	What pre-declared threshold, replication budget, S ₂ plateau, or S ₃ bound justifies commitment?

The ledger serves four purposes: it makes domain-of-validity explicit, tracks where signals hold or fail, prevents silent expansion into unsupported domains, and preserves falsifiability. Ledger entries are append-only records in the UIS sense: once logged, they are revised only through explicit correction/rollback entries, never through silent overwrite.

Appendix B: Minimal Implementation Thresholds

This appendix specifies a practical floor for applying the Structuralization of Empiricism at Level 3 (the divergence-claim level, where the framework is falsifiable). If an application cannot satisfy these minimum requirements, it remains within scope at Level 1 (architectural mapping) or Level 2 (interpretive lens) but cannot generate falsifiable predictions under this framework.

Minimum measurement requirements

To make a Level 3 claim in any domain, at least one of the following must be measurable under controlled conditions.

- **For S_1 :** A quantifiable agreement or variance metric across at least two verified-independent sources (e.g., Cohen's κ , inter-lab variance, meta-analytic heterogeneity I^2 , cross-instrument correlation).
- **For S_2 :** A measurable time-to-resolution or convergence metric under at least two levels of prior constraint strength (e.g., MCMC iterations, reaction time, policy adoption latency, iteration count under uninformative priors).
- **For S_3 :** A measurable response variable under bidirectional constraint variation, with loop area or return-path discrepancy computable (e.g., confidence ratings under evidence accumulation and retraction, threshold shifts in parameter sweeps, policy reversal hysteresis).
- **For update integrity:** A measurable integrity proxy under independently specified improvement or degradation of R, U, or independence conditions. Acceptable proxies include correction lag, time-to-retraction, replication success under verified independence, variance reduction across independent trials, rollback completion rate, or discriminator-satisfaction rate. The intervention must be defined independently of whether stabilization improves; otherwise the test is circular.

Independence audit

Every Level 3 application must include an explicit independence audit answering:

- What sources generated the records?
- Do sources share instrumentation, calibration standards, training data, or institutional incentives?
- What specific steps were taken to verify independence (different labs, different populations, different methods)?
- What residual correlations remain, and how are they bounded (report k_{eff})?

Without this audit, apparent S_1 convergence cannot be distinguished from pseudo-redundancy, and any claimed commitment fails UIS disclosure.

Default metrics

When domain-specific metrics are unavailable, the following defaults provide a starting point:

- **Convergence latency:** Time (in trials, iterations, or calendar units) from initial constraint application to stable selection.
- **Agreement rate:** Proportion of independent observers or runs reaching the same endpoint under shared record access.

- **Hysteresis area:** Integrated discrepancy between forward and reverse constraint-sweep response curves.
- **Correction lag:** Time from confirmed disconfirmation to effective constraint update (retraction, model revision, policy change). See also: Measurement proxies for corrigibility (above).

Scope rule

If you cannot measure at least one S-signature or update-integrity proxy under controlled conditions, the application is out of scope for Level 3 claims. It may still be useful at Level 1 (as a conceptual mapping) or Level 2 (as an interpretive lens), but it cannot generate falsifiable predictions under this framework, and UIS-RS disclosure must declare the level at which the claim operates.

References (selected)

COPE Council. (2025). *COPE Retraction guidelines — English (Version 3)*. Committee on Publication Ethics. <https://doi.org/10.24318/cope.2019.1.4>

Jones, J. C. (2025). *Universal Collapse Theory—Foundations of Collapse* (WP01 v2.0). HoldingLight LLC.

Jones, J. C. (2026a). *Records Across Nature, Life, and Mind: The Persistence Layer of Collapse* (v2.0). HoldingLight LLC. <https://doi.org/10.17605/OSF.IO/7H6DY>

Jones, J. C. (2026b). *The Structuralization of Empiricism: Formalizing the Structural Conditions Under Which Empiricism Stabilizes Knowledge*. HoldingLight LLC. <https://doi.org/10.17605/OSF.IO/J4GZ9>

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and Replicability in Science*. The National Academies Press. <https://doi.org/10.17226/25303>

Nosek, B. A., et al. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>

Swann, W. B., Jr., Jetten, J., Gómez, Á., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review*, 119(3), 441–456. <https://doi.org/10.1037/a0028589>

Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

AI Disclosure. AI tools were used to assist with manuscript preparation. The underlying theory, arguments, and interpretive claims are the author’s own, and the author takes full responsibility for the content.

Citation: Jones, J. C. (2026). Update Integrity Standard (UIS v1.0): A Structural Ethic for Preserving Corrigibility in Record-Based Systems. HoldingLight LLC.

Series: Universal Collapse Theory — Standards Layer

Companions: *Records Across Nature, Life, and Mind; The Structuralization of Empiricism*

© 2026 Jeremy C. Jones — HoldingLight LLC