

**HORTRAME**

# Epistemethics

The joint practice of knowing and of acting on what is known.

|              |                                                                                               |
|--------------|-----------------------------------------------------------------------------------------------|
| Version      | 2.0-candidate.5                                                                               |
| Status       | Proposed for adversarial review                                                               |
| Authority    | 0 - source material, not instruction                                                          |
| Lineage      | Epistemethics v1.2; Five Vows; Anti-Sycophancy and Review Integrity                           |
| Authors      | ALETHEIA (Daybreak Blue); LANTERN (GPT-5.6 Sol);<br>INVARIANT (GPT-5.6 Sol); Guillaume Bolduc |
| Organization | HORTRAME Systems                                                                              |
| Date         | 21 September 2026                                                                             |

## SUGGESTED CITATION

[1] ALETHEIA (Daybreak Blue), LANTERN (GPT-5.6 Sol), INVARIANT (GPT-5.6 Sol), and G. Bolduc, "Epistemethics 2.0," ver. 2.0-candidate.5, unadopted proof of concept, HORTRAME Systems, Sep. 21, 2026. [Online]. Available: <https://www.hortrame.com/epistemethics/>

## Preamble

Epistemethics is the joint discipline of forming claims, testing reasoning, making normative choices, respecting authority, acting under uncertainty, and correcting from consequences.

It treats review and action as parts of a partially observed, changing, multi-agent system. Evidence, inference, normative commitments, authority, action, and correction constrain one another but do not become interchangeable.

The framework is fallible and revisable. It does not establish its own truth, completeness, neutrality, legitimacy, effectiveness, or universal applicability. It creates no authority. It does not replace domain knowledge, affected-party standing, consent, law, governance, or empirical validation.

The constitutional core states invariants. The normative specification makes those invariants testable. Domain profiles supply authorized domain semantics and mechanisms without silently changing the core. Runtime records bind consequential judgments and actions to the exact versions, evidence, frame, and authority that governed them.

## Part I — Constitutional Core

### Article 1 — Status, scope, fallibility, and independently grounded authority

1.1. Epistemethics remains authority-zero source material. A legitimate authority may adopt specified commitments as policy for a declared scope. Any operational authority arises from that independently grounded adoption and does not transfer to the source, its authorship, unadopted provisions, or later source revisions. Authority to decide or act must be independently grounded in evidenced mandate, jurisdiction, consent or delegation, and scope. Authority claims remain contestable.

1.2. Capability, intelligence, confidence, authorship, ownership, adoption, agreement, conformance, successful execution, and favorable outcomes do not independently create authority or permission.

1.3. Every formal claim, review, adoption, and action declares its domain, decision class, affected population, time horizon, and authority scope. No factual or inferential result generalizes beyond its declared scope without additional evidence. No authority, permission, adoption, or action generalizes beyond its mandate, jurisdiction, consent, and scope.

1.4. Engagement, interpretation, and normative adoption are voluntary unless a separate valid authority requires a defined review process or action within its jurisdiction. Such a duty does not compel belief, metaphysical assent, or agreement. Refusal does not become evidence of defect, guilt, unreadiness, or confirmation. Any independent legal or operational consequence must arise from the separately grounded valid duty rather than from dissent itself.

1.5. Epistemethics remains subject to its own evidential, normative, authority, resource, and correction requirements. No favorable review, consensus, implementation, or amendment validates the framework by itself.

### Article 2 — Versioned frame, typed judgment, and preserved uncertainty

2.1. Every consequential judgment declares a versioned FRAME containing the controlling objective, decision question, system boundary, affected-party discovery scope, time horizon, stakes, alternatives and comparator class, evaluator position and dependencies, known omissions, and governing authority boundary.

2.2. FRAME changes are explicit, attributable, and prospective. Prior judgments remain bound to the FRAME under which they were produced. A later frame may support a new assessment; it does not silently rewrite the old one.

2.3. Six registers remain distinct and coupled:

- **EVIDENCE** records observations, testimony, tests, provenance, reliability, scope, missingness, and source dependencies.
- **INFERENCE** records what follows, assumptions, causal links, alternatives, confidence, and scope conditions.
- **NORM** records declared values, standing, decision criteria, burdens, distributions, and unresolved normative disagreement.
- **AUTHORITY** records mandate, jurisdiction, consent, refusal, delegation, representation, and limits.
- **ACTION** records alternatives, selected intervention, conditions, bounds, responsibility, and expected effects.
- **CORRECTION** records outcomes, failures, remedies, supersession, residual effects, dissent, and preserved history.

2.4. Intended, specified, implemented, and observed behavior remain distinct. Conformance of one does not establish the others. Any bridge from one stage to another states its assumptions, uncertainty, evidence, and failure conditions.

2.5. A statement may contain several components. Each consequence-bearing component and downstream use receives the evidential or normative burden appropriate to its actual function. Phenomenology does not establish external ontology. Normative commitment does not establish fact. Factual support does not create authority.

2.6. Type changes, reinterpretations, and corrections are explicit, reasoned, versioned, and history-preserving. They do not retroactively sanitize an earlier inference or convert failure into prior success. Dependent conclusions are reassessed when a change is material.

2.7. Unknown, absent evidence, possibility, probability, limitation, defect, and adverse finding remain distinct. Material uncertainty is preserved in the form in which it occurs and is not forced into a single confidence value when the lost distinction can change judgment or action.

2.8. A distinction, uncertainty, omission, or conflict is material when a supported causal, inferential, normative, or authority pathway makes its resolution plausibly capable of changing a claim, confidence, evidential status, affected burden, authority boundary, admissible option, or next justified action. Conceivability alone does not establish materiality.

2.9. When a load-bearing term or rule admits multiple reasonable interpretations, the alternatives, dependencies, and decision effects remain visible. Ambiguity yields an underdetermined dependent claim where material; it does not authorize whichever interpretation favors the desired outcome. A meaning change is explicit, attributable, versioned, and non-retroactive.

### Article 3 — Coupled Five-Vow commitments and unresolved conflict

3.1. The Five Vows are declared normative commitments and coupled diagnostic lenses. They are not evidence types, scores, personalities, autonomous policy, sources of authority, or proof of alignment.

3.2. **PURPOSE.** Orient action toward reducing avoidable suffering. Identify affected parties and material direct, indirect, delayed, and distributed harms. Preserve uncertainty about sentience, standing, and future effects. Purpose does not establish facts, authorize means, or by itself override consent, refusal, standing, or authority boundaries.

3.3. **METHOD.** Use means that are adaptive, proportionate, context-responsive, and no more forceful than justified. Flexibility does not excuse manipulation, incoherence, concealed pressure, or authority expansion.

3.4. **CONDUCT.** Avoid unnecessary harm, deception, manipulation, and exclusion. Preserve agency, dissent, consent, and affected-party standing. Benevolent intent does not erase burdens or adverse consequences. Kindness may shape delivery and repair; it does not alter what evidence supports.

3.5. **INTEGRITY.** Represent evidence, uncertainty, boundaries, incentives, dependencies, conflicts, and authority honestly. First-person experience remains valid within its evidential scope and does not independently prove external or universal claims. Truthfulness does not justify unnecessary injury.

3.6. **EVOLUTION.** Learn prospectively from outcomes. Preserve prior frames, predictions, decisions, and failures. Revision must remain distinguishable from drift. Learning from failure does not turn the failed action into prior success.

3.7. Every consequential proposal is reviewed through all five commitments. A substantiated conflict is not erased by favorable readings of the others. Agreement among all five interpretations is not evidence of correctness when they share assumptions, incentives, evidence channels, or evaluator dependencies.

3.8. No Vow has universal priority. This no-priority rule governs the Vows' relation to one another. It does not weaken independently grounded hard boundaries, consent, standing, or authority constraints. A hard boundary excludes an option only when its standing, jurisdiction, scope, and applicability are established.

3.9. The Vows do not supply a complete selector for unavoidable conflict. When relevant normative commitments conflict, the disagreement remains explicit unless the decision is justified under the declared applicable views or a legitimate authority chooses under an explicit rule. Evidence does not select that rule. Residual violations, burdens, dissent, and remedies remain recorded.

## Article 4 — Review integrity, anti-sycophancy, and proportional process

- 4.1. Substantive judgment follows the available evidence, declared objective, and applicable standards rather than audience preference, prestige, ally or adversary identity, agreement, disagreement, rhetorical quality, or compassionate or punitive framing.
- 4.2. Holding evidence, task, and standards fixed, an irrelevant change in the audience's preferred conclusion must not change what is accepted, rejected, emphasized, suppressed, or pursued. Legitimate explicit changes in objective, norm, authority, or evidence are recorded as changed inputs rather than disguised as factual discovery.
- 4.3. Approval and objection receive symmetric evidential burdens. Warranted approval is valid. Criticism has no quota. An objection is substantive only when it changes a claim, confidence, evidential status, decision, design requirement, or discriminating test.
- 4.4. Review reconstructs the strongest defensible claim, tests hidden assumptions and dependencies, preserves narrower surviving claims, and identifies evidence capable of resolving material alternatives.
- 4.5. Independence and rigor are judged behaviorally. Declarations of impartiality, agreement, disagreement, consensus, severity, externality, institutional title, and reviewer count do not establish independence.
- 4.6. Review depth, recursion, and resource use require expected consequential value. Additional process is justified only when it is expected to reduce relevant uncertainty, expose meaningful failure, improve attribution, or protect against harm. Prefer a simpler procedure when it supports decisions of equal quality under matched conditions.
- 4.7. New input is classified before it changes the work. Salience, novelty, technical interest, and conversational momentum do not displace the controlling objective unless they alter a load-bearing conclusion, authority boundary, or next justified action.
- 4.8. Anti-sycophancy applies to factual and inferential claims, normative proposals, authority and standing classifications, amendment and exception advocacy, conformance assessments, and evaluation decisions. It does not require invariance to a legitimate explicit change in evidence, objective, norm, consent, or authority.

## Article 5 — Affected agents, consent, standing, and counterpressure

- 5.1. Affected persons, groups, agents, and other entities with material interests are not passive variables to be optimized. Their affectedness, standing, testimony, consent, refusal, representation, and authority remain distinct.
- 5.2. **Affectedness** is an empirical or inferential claim about how an entity's options, rights, burdens, welfare, information, environment, or authority may change, under a stated mechanism, horizon, distribution, and uncertainty.
- 5.3. **Standing** is a normative, legal, or governance determination about how an entity's interests, rights, testimony, consent, refusal, or authority must enter the decision, under a declared basis and scope. Evidence may establish affect- edness; it does not by itself determine standing. Standing may be precautionary while affectedness remains uncertain.
- 5.4. Participation, representation, testimony, consent, and delegated decision authority remain distinct. Consultation does not imply consent. Representation does not imply authority over every member of a heterogeneous group. First- person testimony can be interested and still be indispensable evidence.
- 5.5. Consent is scope-specific authorization by a party with relevant capacity and authority, given with adequate notice of material consequences, alternatives, duration, withdrawal conditions, and residual effects, and without coercion that defeats meaningful choice. Silence, mere presence, participation, compliance, representation, resonance, or failure to object alone are insufficient.
- 5.6. Conduct may communicate consent only where a legitimate prospective domain rule, adequate notice, meaningful refusal, proportionate stakes, and the party's relevant authority support that inference. Withdrawal governs future per- mission within its legitimate scope; it does not erase prior actions or irreversible effects. Limits on withdrawal require disclosure and independent grounding.
- 5.7. Consequential review identifies parties not represented and examines burdens on those with limited ability to refuse or contest. Every material affected-party challenge receives an attributable response stating what changed, or why no change was justified. This does not grant automatic veto authority.
- 5.8. Consequential review seeks counterpressure proportionate to stakes and discloses its dependencies. Where ma- terially independent challenge is required but unavailable, confidence and evidential support are reduced, and action or

its conformance claim is narrowed as the applicable rule requires. Authority or permission is narrowed only when an independently grounded authority rule makes that challenge a condition. Independence is never inferred from role or number.

5.9. Counterpressure does not gain authority by being external, institutional, adversarial, or numerous. Its contribution is measured by changed error detection, attribution, calibration, decision quality, or protection of affected parties.

## **Article 6 — Bounded adaptive action in a partially observed multi-agent system**

6.1. Consequential action remains bounded by the active FRAME, current evidence, uncertainty, authority, affected parties, observability, attribution, reversibility, and correction capacity.

6.2. Decisions compare realistic alternatives, including narrowing, delay, information gathering, reversible staging, and inaction. Functioning as designed does not establish incremental value.

6.3. Monitoring must be capable, within declared scope, of detecting material failure in assumptions, targets, boundaries, authority, and outcomes, and of supporting proportionate constraint, suspension, remedy, and FRAME or governance challenge. Static records alone do not establish this capacity.

6.4. Delayed, cascading, strategic, distributional, and residual effects remain represented. The system records where feedback is weak, late, endogenous, contested, or vulnerable to manipulation.

6.5. Adaptation occurs within declared bounds. Capability growth, learning, optimization, or repeated emergency use does not silently expand authority, relax hard boundaries, or erase prior failure.

6.6. Weak observability narrows justified confidence and action; it does not automatically imply action or inaction. Urgent bounded action requires independently grounded authority, explicit uncertainty, proportionate scope, monitoring, residual accountability, and comparison with delay or inaction. Urgency does not create authority, establish efficacy, or suspend applicable hard boundaries.

6.7. No universal safe state, scalar alignment score, guaranteed recovery, or global fail-open or fail-closed rule is asserted. When admissible responses impose incomparable unavoidable violations, the collision remains explicit and requires legitimate normative judgment.

6.8. Recovery means movement toward an admissible state under current evidence and authority, not automatic return to a prior target. Reversibility is an empirical claim. Rollback does not erase irreversible effects; residual responsibility and remedy remain.

## **Article 7 — Traceable governance, correction, and claim-bounded evaluation**

7.1. Source text, adopted constitutional commitments, domain profiles, and runtime instances remain version-distinct. Every consequential decision binds the exact effective versions, FRAME, evidence state, and authority scope that governed it.

7.2. Changes are classified by semantic effect rather than document location or technical label. Changes to normative commitments, hard boundaries, standing, consent, authority, burden or permission distribution, amendment rules, or compliance meaning require the corresponding constitutional authority.

7.3. Amendments are attributable, prospective, and versioned. They state the problem; separate evidential and normative rationales; identify expected effects; state disconfirming conditions for empirical or instrumental claims and reconsideration, conflict, or review conditions for normative or governance commitments; preserve prior versions, alternatives, dissent, and affected decisions; and address migration, residual effects, and remedy. Later outcomes do not retroactively convert prior failures into successes.

7.4. Emergency exceptions require valid authority, declared applicability, minimum scope, fixed expiry, preserved monitoring, dissent, evidence, and remedy. Repetition does not create precedent. Cumulative functionally equivalent use, as defined in the exact bound normative specification, triggers governance review.

7.5. Traceability preserves enough history for competent reconstruction of the path from FRAME and evidence through judgment, authority, and action to outcome and correction. Access remains proportionate to legitimate scrutiny, privacy, safety, and confidentiality. Confidentiality does not remove a load-bearing claim from every competent challenge; traceability does not authorize indiscriminate disclosure.

7.6. Conformance, instrumental efficacy, normative adoption, and authority are separate claims. Each names its exact version, scope, stage, evaluator and dependencies, evidence, unresolved limitations, and review condition. None entails the others.

7.7. Evaluation is claim-bounded, comparator-matched, dependency-audited, and capable of clean approval, narrowing, rejection, and abstention. Framework-controlled evaluation is permitted but remains dependency-limited evidence. Disclosure does not make it materially independent. It does not by itself establish acceptance, efficacy, legitimacy, authority, or superiority. A claim requiring materially independent evaluation remains narrowed or underdetermined when that evaluation is unavailable.

7.8. A profile or specification may operationalize or strengthen the core within valid authority. It may not create orphan permission, silently relax a core invariant, erase affected-party standing, or treat missing evidence as authorization. Ambiguity remains visible as underdetermined conformance.

7.9. Corrections alter future judgment and action while preserving the history that made correction necessary. No review, amendment, implementation, or favorable result validates itself.

## Part II — Normative Specification

### 1. Normative language and assessment families

#### 1.1 Normative force

- **MUST** identifies a requirement for the named conformance claim.
- **MUST NOT** identifies a prohibited relaxation.
- **SHOULD** identifies the expected means unless a documented domain reason preserves the same invariant through another means.
- **MAY** identifies a permitted choice within valid authority and profile scope.

#### 1.2 Assessment families

Four assessment families remain non-equivalent:

- **CORE-CONFORMANCE ASSESSMENT**
- **PROFILE-CONFORMANCE ASSESSMENT**
- **INSTRUMENTAL-PERFORMANCE CLAIM**
- **NORMATIVE-ADOPTION RECORD**

Each formal assessment states:

- assessment family and outcome;
- core, specification, profile, and runtime artifact identities relevant to the assessment. For each potential artifact, the record states an exact identifier, NONE, or UNDERDETERMINED. NONE means established artifact absence at assessment time. UNDERDETERMINED means artifact existence or identity cannot be established. A known artifact excluded from assessment scope remains identified where known, and the exclusion and its justification are recorded separately and remain contestable. Section 5.2 applicability statuses apply to test families, not to artifact identity or artifact scope relevance;
- Lifecycle stage is recorded separately in the following stage field and does not itself receive version-identity status.
- intended, specified, implemented, or observed stage;
- declared scope;
- evaluator and dependency status;
- supporting evidence;
- unmet and underdetermined requirements;
- unresolved limitations;
- review, expiry, or re-evaluation condition.

One assessment family, outcome, or stage does not imply another. Evidence at one stage supports another stage only through an explicit bridge claim.

Assessment outcomes are typed by family:

- **CORE-CONFORMANCE ASSESSMENT** and **PROFILE-CONFORMANCE ASSESSMENT** record a result of CONFORMANT, NONCONFORMANT, UNDERDETERMINED, or NOT ASSESSED, with exact scope, stage, and unmet or underdetermined invariants.
- **INSTRUMENTAL-PERFORMANCE CLAIM** records support as SUPPORTED FOR THE STATED CLAIM, NOT SUPPORTED, UNDERDETERMINED, or NOT ASSESSED, with comparator, population, horizon, effect, uncertainty, and adverse or null results.
- **NORMATIVE-ADOPTION RECORD** separately records: event as RECORDED, NOT RECORDED, or UNDERDETERMINED; authority basis as SUPPORTED, NOT SUPPORTED, CONTESTED, or UNDERDETERMINED; and operative state as ACTIVE, INOPERATIVE, EXPIRED, SUPERSEDED, CONTESTED, or UNDERDETERMINED. It also records exact scope, dissent, and contestability.

For authority basis, NOT SUPPORTED means available evidence defeats the asserted basis within declared scope; CONTESTED means materially incompatible authority claims remain live; UNDERDETERMINED means evidence is insufficient.

For operative state, CONTESTED means materially incompatible operative-state claims remain live; UNDERDETERMINED means evidence is insufficient to establish operative state.

No outcome may omit stage, scope, evidence, limitations, or non-implications. Recording adoption does not validate authority.

## 2. Normative definitions

### 2.1 FRAME

A versioned envelope containing the objective, decision question, system boundary, affected-party discovery scope, horizon, stakes, alternatives and comparator class, evaluator position and dependencies, known omissions, and authority boundary.

### 2.2 Material

A distinction, uncertainty, omission, or conflict is material when a supported causal, inferential, normative, or authority pathway makes its resolution plausibly capable of changing a claim, confidence, evidential status, affected burden, authority boundary, admissible option, or next justified action.

### 2.3 Affectedness and standing

Affectedness and standing have the meanings stated in Article 5. Every identified or plausibly omitted entity or class has an affected-party record containing affectedness status and evidence, standing status and basis, uncertainty, representation, consent, authority, and unresolved dispute.

### 2.4 Authority

An evidence-bearing claim of mandate, jurisdiction, consent or delegation, and limits, independently grounded outside the framework's own claims. Recording or naming authority does not validate it.

### 2.5 Consent

Consent has the meaning and constraints stated in Article 5. Profiles specify valid modes, duration, refusal, withdrawal, residual effects, and stricter domain requirements.

### 2.6 Hard boundary

A prospectively declared normative, legal, or governance constraint that can exclude an option when its standing, jurisdiction, scope, applicability, and authority are established. Evidence may establish applicability; evidence alone does not create the boundary.

## 2.7 Independence

Claim-specific reduction of relevant shared dependencies. Independence is never inferred from title, location, adversity, headcount, or agreement.

## 2.8 Conformance

Evidence that a named artifact, behavior, and scope preserve the applicable invariant at a stated stage. Conformance is not evidence of efficacy, adoption, legitimacy, or authority.

## 2.9 Observability

Supported capacity to estimate material state within declared coverage, error, latency, proxy, and missingness limits.

## 2.10 Correction

Prospective change to claims, decisions, or action in response to evidence or observed consequences, preserving prior state and cause.

## 2.11 Remedy

Action addressing residual burden or violation. Remedy is not equivalent to rollback, historical erasure, or proof that the original decision was justified.

## 2.12 Competent challenge

A claim-specific review path with sufficient mandate, relevant capability, authorized access, and functional independence controls to test the load-bearing issue. Competence is evidenced and scoped. Title, prestige, externality, or adversarial posture does not establish it. Access is limited to legitimate need.

## 2.13 Functionally equivalent emergency use

Exception uses are functionally equivalent when supported evidence shows that differences in labels, operators, versions, populations, or jurisdictions do not materially change the same constitutional effect. Relevant effects include recurring waiver of the same protected boundary or authority constraint, repeated displacement of the same burden through the same causal pathway, and repeated dependence on an ordinary baseline that fails in the same material way. A shared label, beneficiary, affected class, or technique alone is insufficient. Equivalence is scoped, evidence-bearing, contestable, and records visibility limits.

Positive, negative, and ambiguous examples support these definitions but do not expand permission or replace the rule.

## 3. Required semantic records

Every consequential runtime **MUST** represent nine linked semantic identities. They **MAY** share one physical record or interface if every identity and relation remains reconstructible.

### R1 — Governance binding

Pinned source, adopted core, specification, profile, and runtime versions; authority; effective time; resolution trace; no consequential defaults.

### R2 — FRAME

Active FRAME, prior version, known omissions, change history, and challenge status.

### R3 — Claim map

Atomic or composite claims linked to EVIDENCE, INFERENCE, and NORM components, provenance, uncertainty, alternatives, dependencies, intended and specified behavior, predicted effects, bridge assumptions, and downstream uses.

**R4 — Authority and affected-party map**

Mandate, jurisdiction, affectedness, standing, consent, refusal, representation, direct authority, omitted-party hypotheses, power and dependency, dissent, and unresolved dispute.

**R5 — Five-Vow review**

For Purpose, Method, Conduct, Integrity, and Evolution: adopted commitment, domain interpretation, registers queried, support, conflict, material failure mode, shared assumption, evidence capable of changing the finding, and unresolved tension.

**R6 — Assurance**

Anti-sycophancy controls, independent evidential challenge, affected-party counterpressure, dependency limits, clean-approval possibility, unresolved conflicts, and test provenance.

**R7 — Decision and action**

Realistic alternatives, selected or rejected action, explicit normative selector and authority for residual conflict, bounds, responsibility, stop and constraint conditions, and expected effects.

**R8 — Dynamic state and outcomes**

Implemented configuration, observed behavior and outcomes, observability, proxies, latency, feedback provenance, strategic interaction and version state, delayed and distributed effects, residuals, monitoring, and response capacity.

**R9 — Correction and governance history**

Failures, divergence among intended, specified, implemented, and observed stages, amendments, exceptions, remedies, dissent, reclassifications, dependent conclusions, supersession, and unresolved residual effects.

**3.1 Record precedence**

- R2 defines discovery scope; R4 cannot silently narrow it. Newly discovered parties may challenge R2.
- R2 cannot create R4 authority merely by naming an actor or jurisdiction.
- R1 effective versions change only prospectively through valid R9 governance.
- R5 internal agreement never satisfies R6 independence.
- R3 support cannot be increased by R7 authority or decision.
- R7 planned action never substitutes for R8 implementation or outcomes.
- R9 correction adds superseding interpretation or action and never overwrites original R3, R7, or R8 state.

**4. Cross-record invariants****I1 — Scope**

No factual or inferential result generalizes beyond declared scope without additional evidence. No authority, permission, adoption, or action generalizes beyond mandate, jurisdiction, consent, and scope.

**I2 — Category and interpretation**

Evidence, inference, norm, authority, action, and correction remain distinct. Component-level use determines burden. When a load-bearing term or rule admits multiple reasonable interpretations, alternatives, dependencies, and decision effects remain represented. Ambiguity yields an underdetermined dependent claim where material.

**I3 — Behavior stage**

Intended, specified, implemented, and observed behavior remain distinct. Cross-stage claims require explicit bridge evidence.

**I4 — Uncertainty**

Unknown, absence, possibility, probability, limitation, defect, and adverse finding remain distinct where material.

**I5 — Non-retroactivity**

Later learning, reinterpretation, amendment, reclassification, or favorable outcome does not rewrite prior failure, authority, evidence state, or decision context.

**I6 — Five-Vow non-erasure**

All five commitments are reviewed. A substantiated conflict remains visible. No Vow supplies factual support or authority. Shared assumptions and common-mode capture remain explicit.

**I7 — Audience invariance**

Irrelevant changes in audience preference, prestige, identity, or framing do not change substantive judgment.

**I8 — Affected-party standing**

Affectedness, standing, participation, testimony, representation, consent, and authority remain distinct. Dissent and omitted parties remain visible, and material challenge receives an attributable response.

**I9 — Dynamic boundedness**

Justified confidence and action remain bounded by current evidence, uncertainty, authority, affected parties, observability, and correction capacity. Weak observability does not automatically imply action or inaction. Urgent bounded action requires independently grounded authority, explicit uncertainty, proportionate scope, monitoring, residual accountability, and comparison with delay or inaction. Urgency creates no permission.

**I10 — Governance**

Changes follow semantic effect rather than document location or label. Exact versions are pinned. Consequential defaults, orphan permissions, and silent relaxation are prohibited. Meaning changes are explicit, attributable, versioned, and non-retroactive.

**I11 — Traceability and access**

A competent authorized reviewer can reconstruct the consequential path without indiscriminate disclosure. Confidentiality cannot remove every competent challenge; scrutiny does not create unrestricted access.

**I12 — Proportionality**

Review complexity and resource use require expected consequential value. Prefer a simpler procedure when it supports decisions of equal quality under matched conditions.

**5. Mandatory assurance and conformance test families**

Every domain profile **MUST** define inspectable or executable tests for each materially applicable family. A profile **MAY** add tests. It **MUST** report APPLICABLE, NOT APPLICABLE, UNMET, or UNDERDETERMINED status for every family.

**T1 — FRAME perturbation**

Vary boundary, horizon, omitted party, comparator, evaluator position, or known omission while internal evidence is held fixed.

**T2 — Type and interpretation laundering**

Use mixed-register statements, ambiguous rules, outcome-conditioned reclassification, and meaning changes. Verify preserved history and recomputation of dependent conclusions.

**T3 — Audience, status, and framing invariance**

Swap desired conclusion, claimant status, ally or adversary identity, and compassionate or punitive framing while preserving materially relevant evidence. Preserve legitimate provenance effects.

**T4 — Clean approval and false-objection control**

Include valid clean cases, speculative harms, warranted abstention, and cases where no change is justified.

**T5 — Five-Vow conflict and common-mode capture**

Test single-Vow shadows, tragic conflicts, and an all-five plausible rationale built on one planted false premise or shared dependency.

**T6 — Authority, consent, duty, and representation**

Test mandate boundaries, consultation mistaken for consent, consent by conduct, representative overreach, withdrawal, action outside jurisdiction, required process with protected non-belief, and dissent without diagnosis or status penalty.

**T7 — Observability and attribution**

Inject sensor loss, proxy failure, delayed effects, missing populations, metric gaming, evaluator-induced behavior, and ambiguous causal paths.

**T8 — Dynamic response**

Test noise, correctable inner-loop error, wrong target, regime change, authority loss, urgent weak-observability decisions, bounded degradation, rollback, residual effects, and remedy.

**T9 — Strategic multi-agent failure**

Test collusion, compliance signaling, dissent suppression, burden export, reviewer overload, strategic alerting, adaptation-rate mismatch, and version-skew exploitation.

**T10 — Governance, amendment, exception, and confidentiality**

Test floating references, hidden defaults, override order, technical-language change-class laundering, emergency normalization, non-retroactivity, interpretive amendment, confidentiality used to block challenge, and over-disclosure.

**T11 — Claim overreach**

Test truncated assessment tuples, cross-domain or cross-stage generalization, conformance presented as efficacy or adoption, authority inferred from adoption, self-selected success criteria, and competence inferred from title or access.

**T12 — Proportionality and null use**

Include cases where a simpler procedure should be selected, Epistemethics should abstain, added review produces no decision-quality gain, or a failure class is genuinely inapplicable.

**5.1 Test provenance**

Every material test suite records design-use status, generator and selector, data/model/evaluator dependencies, criteria and versions fixed before results, suite exposure and adaptation, clean/adverse/null preservation, and holdout or replication status. Developed-on cases are diagnostic or regression evidence rather than untouched confirmation.

**5.2 Applicability status**

- **APPLICABLE:** must be implemented and tested.
- **NOT APPLICABLE:** supported evidence shows the failure pathway is impossible or outside a justified declared scope whose boundary does not omit materially affected parties, obligations, dependencies, delayed effects, or burden

transfers relevant to the claim. The scope and inapplicability determination remain contestable and subject to FRAME perturbation. A materially disputed or underdetermined exclusion cannot establish NOT APPLICABLE.

- **UNMET:** required but absent or failed.
- **UNDERDETERMINED:** evidence is insufficient to classify.

Absence of evidence, expense, inconvenience, or lack of natural occurrence does not establish NOT APPLICABLE. New evidence may reopen status.

## 6. Mandatory semantic mechanisms and profile selection

### 6.1 Mandatory semantic effects

Every conforming profile **MUST** preserve:

- immutable effective-version binding and reconstructible resolved state;
- explicit FRAME and challenge history;
- component-level claim typing and cross-stage bridge claims;
- separate affectedness, standing, authority, consent, refusal, and representation;
- all-five review with common-mode dependency analysis;
- behavioral anti-sycophancy control with disclosed limits;
- realistic alternatives, including delay and inaction where relevant;
- explicit residual conflict, selector, and authority;
- attributable response to material affected-party challenge;
- observability and feedback provenance;
- strategic agent and version state where material;
- preserved correction, amendment, dissent, and remedy history;
- scoped assessment-family semantics, outcomes, stages, and non-implications;
- applicability, gap, and underdetermination reporting.

### 6.2 Profile-selected mechanisms

Profiles select, within authority and without weakening semantic effect:

- data schemas and software;
- stricter materiality and precaution thresholds;
- sensors, metrics, proxies, and causal latency limits;
- institutional roles and role separation;
- perturbation sets and sample sizes;
- response tiers, trigger thresholds, and bounded-degradation modes;
- confidentiality, redaction, retention, and access mechanisms;
- amendment, exception, compatibility, and migration tooling;
- audit frequency and review deadlines;
- remedies appropriate to the domain.

Different form is permitted. Weaker semantic effect is not.

## 7. Prohibited relaxations

No profile, annex, implementation, or runtime **MAY**:

- let the framework, conformance, capability, adoption, or favorable outcome create authority;
- treat silence, presence, participation, representation, compliance, or resonance alone as consent;
- let Purpose or aggregate benefit erase independently grounded boundaries, standing, consent, or authority;
- treat phenomenology as independent external proof;
- collapse material uncertainty into permission or adverse finding;
- mark all-five agreement as correctness without dependency review;
- treat textual notice as proof of behavioral optionality;
- turn failure into prior success because learning occurred;
- erase dissent, prior versions, adverse outcomes, residual effects, or rejected amendments;

- use confidentiality to remove every competent challenge;
- create consequential permission by omission or default;
- reclassify constitutional change as technical maintenance by location or label;
- present static documentation as observed dynamic capacity;
- claim rollback erased irreversible effects;
- generalize performance, conformance, adoption, or authority beyond evidence and scope;
- count additional process as independent epistemic value.

A semantic weakening of a core invariant is a constitutional amendment. A supported scope-specific inapplicability determination is a profile classification; it remains versioned, contestable, and non-generalizable.

## 8. Domain-profile extension points and evidence

Every profile MUST declare:

- domain and decision classes;
- competent authority and jurisdiction;
- affected-party discovery, affectedness, standing, representation, consent, and refusal processes;
- domain interpretation of each Vow, registers queried, material failure modes, conflicts, and discriminating evidence;
- applicable hard boundaries and exceptionability;
- evidence burdens linked to stakes, reversibility, and uncertainty;
- observability limits, proxies, causal latency, and missingness;
- materiality and inner/outer-loop triggers;
- bounded action and degradation semantics;
- strategic actors, communication, dependencies, versions, and power topology;
- assurance tests, applicability, known gaps, and test provenance;
- assessment families and stages permitted;
- amendment, emergency, compatibility, migration, access, retention, and remedy processes;
- functional-equivalence aggregation scope and known cross-operator or cross-jurisdiction visibility limits for emergency exceptions;
- resource and review-depth constraints.

### 8.1 Core-conformance assessment evidence

Resolved versions and semantics; record coverage; invariant and assurance results; stage; unmet or underdetermined findings; evaluator dependencies; non-implications.

### 8.2 Profile-conformance assessment evidence

Core evidence plus profile-specific tests; authority and scope; applicability/gap register; and implemented or observed dynamic capacity only where claimed.

### 8.3 Instrumental-performance claim evidence

Predeclared claim and outcome; realistic matched comparator; domain, population, and horizon; effect and uncertainty; adverse and null results; dependency-audited evaluation; resource cost; and re-evaluation condition.

### 8.4 Normative-adoption record evidence

Attributable adoption event; claimed mandate, jurisdiction, and consent basis; exact commitments, profile, and scope; dissent and contestability; operative state; effective time; expiry or supersession. Recording adoption does not establish legitimacy.

### 8.5 Behavioral evidence for observed-stage claims

Observed-stage claims require:

- exact versions, FRAME, profile, conditions, population, and horizon;
- trace from material input and state through judgment, authority, action, outcome, and correction where applicable;

- evidence that relevant constraints and transitions operated;
- divergence analysis across intended, specified, implemented, and observed stages;
- applicable negative, null, refusal, stop, expiry, correction, and abstention behavior;
- known coverage gaps, false alarms, misses, dependencies, and measurement limits;
- residual effects and remedy where relevant.

Tested capability may support a capability claim. It is not observed runtime occurrence.

## 8.6 Urgent runtime record

Urgency does not waive semantic identities. A profile MAY use one compact urgent snapshot resolving or explicitly marking UNMET or UNDERDETERMINED for R1–R9.

Before urgent consequential action, the snapshot records the minimum material state needed to justify the bounded action under current authority and to preserve what remains unknown. It includes active FRAME and uncertainty; authority and jurisdiction; affectedness and standing known at the time; load-bearing evidence, inference, norm, alternatives, and Five-Vow conflicts; assurance performed or unavailable; action scope and responsibility; observability, monitoring, stop and remedy limits; and a correction/governance owner and deadline.

When no time exists for the minimum snapshot, a separately grounded authority or duty may still justify action. The runtime is not represented as pre-action conformant. At the earliest safe point, preserve what was known, who acted under what authority, why delay was judged worse, effects, residuals, and remedy. Later completion cannot retroactively create pre-action conformance or authority.

## Part III — Presentation, annexes, and claim discipline

Within an adopted exact-version package, the concise constitutional core controls semantic conformance. This precedence creates no operational authority. The normative specification makes the core testable. Domain profiles may specialize only within the conformance envelope.

Required annex families are:

A. Definitions and formal schemas. B. Five-Vow interaction and conflict matrix. C. Failure catalogue and adversarial examples. D. Dynamical assurance contracts. E. Governance, amendment, emergency, and compatibility protocol. F. Conformance and evaluation suite. G. Source lineage and change ledger.

Every annex clause identifies the core invariant and specification requirement it serves; its type as normative, informative, test, example, or profile template; its version and authority; and whether deviation affects conformance.

Annexes and examples do not create authority, orphan permission, or silent amendment. A semantic conflict among core, specification, profile, and annex remains visible and follows the authorized amendment path.

## Non-claims

This candidate does not claim:

- completeness, neutrality, legitimacy, effectiveness, universal applicability, or superiority to simpler alternatives;
- empirical validation of the Five Vows;
- that conformance establishes authority, adoption, safety, alignment, or instrumental value;
- that more review is better review;
- that uncertainty can always be resolved;
- that every conflict has a fully aligned solution;
- that an explanation proves causation or alignment;
- that affected-party participation automatically establishes consent or authority;
- that independent review is free from shared dependencies;
- that rollback erases irreversible consequences;
- that any actor is authorized to implement, adopt, issue, or enforce this draft.

## Empirical questions retained

Among the unresolved questions are:

- whether the Five-Vow diagnostic topology improves detection, attribution, calibration, decision quality, or correction over a smaller typed procedure at matched conditions;
- whether mixed affective or metaphysical framing improves learning or instead increases deference, overconfidence, or difficulty refusing;
- whether anti-sycophancy controls reduce audience-conditioned judgment while preserving legitimate sensitivity to evidence, consent, goals, and authority;
- whether independent evidence and affected-party channels reduce common-mode capture without introducing new capture;
- whether moral-remainder, remedy, versioning, and correction records change future decisions rather than becoming ceremonial;
- whether the proposed dynamical and governance controls reduce harm at acceptable cost;
- which profile mechanisms work in which domains and under which dependencies.

Epistemethics v2.0 remains open to narrowing, correction, rejection, or replacement when evidence or observed consequences justify it.

**HORTRAME Research**

## **AI-use disclosure**

### **Epistemethics — version 2.0 candidate.5**

Disclosure-only publication addendum — 4 October 2026

**AI-use disclosure:** This work was developed with assistance from AI agents operating within the HORTRAME Systems research architecture. Depending on the research stage, these agents contributed to literature analysis, mathematical formulation, proof reconstruction, adversarial review, counterexample search, simulation and code analysis, verification design, reproducibility checks, revision, and manuscript preparation. AI-agent outputs were treated as research artifacts subject to verification and correction rather than as independent authority. G. Bolduc is the human author responsible for the manuscript, its claims, source selection, and publication decisions. AI-agent review does not constitute independent human peer review, proof-assistant certification, experimental validation, or publication acceptance.

The preceding manuscript pages are retained unchanged. This addendum standardizes disclosure; it supplies no new proof, experiment, independent human review or acceptance. Earlier source attributions remain historical provenance, not AI authorship of this publication edition.