From Episodes to Primitives: Causal Historical Assets for Evidence-Governed Ontology Revision

Keren Dong
Kungfu Origin Technology Limited
keren.dong@kungfu.link

Alpha draft: July 2026

Abstract

Agent systems usually preserve histories to debug prior actions or retrieve prior answers. This paper studies a stronger possibility: a well-formed causal EPISODE corpus may preserve the option to discover that the ontology under which earlier questions were asked was incomplete. We define an Episode as a perspective-bound, provenance-carrying record of observations, choices, actions, dependencies, consequences, interventions, omissions, and adjacent fact cuts. We then specify a staged mechanism derived from Kung Fu Decisions (KFD): KFD-4 replays causal histories from declared perspectives; KFD-5 separates candidate genesis from qualification; draft KFD-6 compares plural generation methods against fixed-ontology and no-new-primitive baselines, held-out evidence, and separated promotion authority. The resulting asset is not an event pile, transcript, embedding index, or automatic proof. Its future value depends on integrity, perspective, causal and responsibility boundaries, portability, retention, privacy, consent, and independent verification. We ground the design in one immutable July 2026 KFD evidence cut containing three live cases and six candidate tracks: three accepted only within software-domain profiles and three provisional. This is early feasibility evidence, not a conforming KFD-6 experiment. We contribute an artifact taxonomy, an evidence-governed discovery loop, a governance analysis of problem-definition power, and a falsifiable prospective evaluation protocol. The central claim is conditional: governed causal history can preserve future ontology-revision capacity; storage volume, narrative coherence, or generator confidence cannot.

1 Introduction

Most computational histories are collected to answer questions already known at capture time: which call failed, which state changed, which action produced a reward, or which artifact was released. The schema names the objects in advance. Later work searches, aggregates, or predicts within that object world.

Real work creates a harder failure mode. A team can preserve every field it knows and still lose the relation that a future participant needs, because the relation did not yet have a name. Permission can be mistaken for accepted responsibility. An executed action can be mistaken for a settled obligation. A collection of verified components can be mistaken for the one project state a successor is authorized to continue. The missing information is not always a missing value. It may be a missing object boundary.

This paper asks whether historical infrastructure can preserve a capacity to repair that kind of omission. The proposal builds on three Kung Fu Decisions: KFD-4 binds views to declared perspectives and permits source-preserving replay and transformation; KFD-5 separates candidate

genesis from fact-bound qualification; draft KFD-6 proposes a bounded loop that compares generation methods over causal experience without allowing a generator to certify or promote itself [6, 7, 8].

The central thesis is conditional:

> A well-formed causal Episode corpus can preserve more than prior answers. It can preserve a governed option to replay decisions, expose unnamed burdens, qualify candidate Primitives, and revise the ontology under which future questions are formed.

The word *option* matters. A corpus does not discover merely by existing. A generated label is not a Primitive. A plausible retrospective story is not causal evidence. The proposed value appears only when later participants can identify an immutable evidence cut, reconstruct perspectives with visible loss, compare plural methods and conservative baselines, reject attractive false candidates, and keep promotion authority separate from generation.

The contribution is distinct from two adjacent KFD papers. The Foundation paper develops facts, trust, cooperation, and boundary primitives as an ordered systems model [2]. The Observer paper develops declared timelines and perspective-preserving replay [3]. This paper treats replayable Episodes as a long-lived input to ontology revision and focuses on the KFD-4 to KFD-5 to KFD-6 transition.

We make five contributions:

1. We define an EPISODE historical asset and distinguish it from a transcript, event log, provenance graph, embedding index, and endpoint state.
2. We specify a staged discovery mechanism that keeps perspective replay, candidate genesis, qualification, held-out evaluation, and promotion authority independently inspectable.
3. We explain why control of historical representation affects the power to define future questions, while deriving stewardship obligations rather than a command to retain everything.
4. We freeze one implementation-backed KFD evidence cut and grade what it supports: three live cases and six tracks, with three bounded software-domain acceptances and three provisional candidates.
5. We propose prospective ablation, shadow-loop, transfer, and governance evaluations capable of falsifying the stronger thesis.

The paper does not claim autonomous Primitive discovery, universal ontology, historical novelty of current KFD candidates, or a quantified return on Episode retention. It claims that a mechanism can be specified, its early feasibility can be separated from its missing gates, and the next experiment can be made answerable to evidence.

2 Problem Statement and Research Questions

Let an operational system at time n act with ontology O_n. Its capture schema determines which entities, relations, state changes, and responsibilities are directly representable. A fixed-ontology history $H(O_n)$ can be perfectly internally consistent while remaining unable to express a distinction x that becomes load-bearing later.

An ontology revision is justified only when the new distinction changes a required decision and cannot be recovered from the old representation at lower total burden. The problem is

therefore not to maximize labels. It is to detect when repeated reconstruction, hidden coordination, divergent consequences, or causal variables indicate that O_n is missing an independently addressable responsibility.

We study four research questions:

RQ1: Representation.
Which properties distinguish an Episode corpus that preserves future ontology-revision capacity from anonymous traces or endpoint facts?

RQ2: Mechanism.
Can declared perspective replay improve candidate genesis while KFD-5 qualification prevents narrative or method artifacts from becoming Primitives?

RQ3: Evaluation.
Under shared evidence and resource budgets, does an Episode-based loop outperform fixed-ontology and no-new-Primitive baselines on held-out reconstruction, transfer, and false-candidate rejection?

RQ4: Governance.
Which integrity, retention, deletion, consent, portability, access, and promotion boundaries are necessary when a corpus may influence not only answers but future problem definitions?

These questions deliberately separate technical possibility from legitimate authority. A system may generate a useful candidate and still lack the right to retain the underlying experience, represent a participant, or change an adopted ontology.

3 Episode Historical Assets

3.1 From recorded events to causal experience

Event order is necessary but insufficient. Distributed histories already show that order and causality cannot be inferred from a single wall clock [11]. Provenance standards distinguish entities, activities, agents, derivation, and responsibility [12]. KFD adds a participant and ontology boundary: what did an observer accept, which action was attempted under which authority, what consequence was borne, what evidence was missing, and which later interpretation remains a claim?

We model an Episode e as:

$$e = (i, P, F^-, A, D, C, I, F^+, R, G, S), \tag{1}$$

where i is stable identity; P is declared observer and perspective; F^- and F^+ are adjacent admitted Fact cuts; A is the action and bounded intent; D records dependencies and authority; C records consequences, costs, failures, retries, and overrides; I records interventions; R binds evidence and receipts; G declares gaps, redaction, degradation, and proxy-only consequences; and S declares stewardship policy.

This tuple is a responsibility decomposition, not a required physical layout. An implementation may use several stores or one content-addressed bundle. It qualifies as an Episode only when the relevant distinctions remain recoverable and verifiable.

3.2 Artifact taxonomy

Table 1 prevents a common category error: more storage does not imply a stronger historical asset.

Artifact	Useful capability	Insufficient by itself because
Transcript	Recover utterances and tool output	Speaker position, action authority, omitted evidence, consequences, and admission may be implicit.
Event log	Query ordered or timestamped events	The event schema may flatten observer, responsibility, causal dependency, and competing views.
Provenance graph	Relate entities, activities, agents, and derivations	It does not necessarily preserve the ontology under test, qualification gate, or promotion authority.
Endpoint Fact cut	State what was admitted at one boundary	It can omit the path, alternatives, intervention, failure, and consequence that explain why a new distinction emerged.
Derived index	Retrieve summaries, embeddings, or clusters efficiently	It is a projection whose loss, source cut, update policy, and non-authority must remain explicit.
Episode asset	Replay a perspective-bound causal occurrence with evidence and governance	It preserves discovery options but neither generates nor qualifies a Primitive automatically.

Table 1: Historical artifact taxonomy. The categories may compose; they are not interchangeable evidence grades.

3.3 Asset conditions

A corpus $\mathcal{E}_c$ is an Episode historical asset only if its declared cut c satisfies six conditions:

1. **Integrity:** stable identities, roots, corruption detection, recovery boundaries, and immutable prior cuts.
2. **Perspective:** observer position, accepted facts, natural objects, consequences, and known gaps remain visible.
3. **Causal and responsibility structure:** actions, dependencies, authority, outcomes, interventions, and omissions are not flattened into anonymous tokens.
4. **Portability:** export and import preserve exact roots or declare loss; provider migration does not silently redefine identity.
5. **Governance:** retention, deletion, consent, access, redaction, purpose, and review policies are enforceable and auditable.
6. **Independent challenge:** another accountable participant or verifier can inspect evidence, degradation, and later qualification.

The corpus may be intentionally incomplete. Completeness claims are stronger than asset claims and usually unavailable. A useful asset makes omissions inspectable instead of generating certainty to cover them.

3.4 Option value without storage determinism

Let Q_f be a future distribution of admissible questions and Π the policies that permit a participant to inspect a corpus. We define the *discovery option value* of an Episode cut qualitatively as the

avoidable future burden of reconstructing perspective, causality, and responsibility when a question requires concepts absent from O_n:

$$V_{\text{option}}(\mathcal{E}_c) = \mathbb{E}_{q\sim Q_f}\left[B(q \mid H(O_n)) - B(q \mid \mathcal{E}_c, \Pi)\right]. \tag{2}$$

B may include reconstruction time, coordination load, missing-evidence risk, or inability to test an ontology revision. The expression is not a monetary valuation and does not imply $V_{\text{option}} > 0$. Poor capture, restricted access, privacy harm, corrupted roots, or ontology lock-in can make retention neutral or harmful. Storage volume is absent from the definition because volume does not establish recoverability or legitimacy.

4 The KFD-4/5/6 Discovery Mechanism

The mechanism is a sequence of gates, not one model call. Each stage preserves a different authority boundary.

4.1 KFD-4: replay changes the observation horizon

For admitted facts F, a perspective P defines a projection $\pi_P(F)$. A transformation T_{PQ} reconstructs a new declared view without claiming to become the source observer:

$$T_{PQ} : (P, \pi_P(\gamma)) \to (Q, \pi_Q(\gamma), J, L), \tag{3}$$

where γ is a causal path, J is the set of preserved invariants, and L is declared loss. Source identity, Fact and evidence cuts, known causal relations, receipts, and authority boundaries may belong to J. Relevance, ordering, natural object boundaries, and available action may change.

Contrastive replay keeps at least two source views distinct. Its purpose is not to vote on one universal view. It asks which burden becomes natural for the participant who bears the next consequence. A user may expose an object hidden from an implementation view; a successor agent may expose a settlement object hidden from the coordinator who assembled its components.

KFD-4 is one genesis method. Direct situated judgment, anomaly search, reconstruction analysis, causal-variable discovery, and structural compression remain first class. This method plurality prevents a successful perspective story from becoming a ranking theorem.

4.2 KFD-5: generation is not qualification

Genesis binds the perspective, current ontology, evidence cut, method, observation, candidate, and claim boundary. Qualification then asks whether the candidate deserves independent identity, boundary, authority, lifecycle, or operations.

For candidate x, define $\text{remove}_x(h)$ as a history that erases the semantic distinction x while preserving other observable facts. The candidate is non-redundant for decision d when two valid histories exist such that:

$$\text{remove}_x(h_1) = \text{remove}_x(h_2) \quad \wedge \quad D_d(h_1) \neq D_d(h_2). \tag{4}$$

This deletion witness does not privilege a name or physical representation. A rival model passes if it preserves equivalent information with lower total burden. KFD-5 therefore requires alternatives, minimum closure, deletion and fuse tests, falsifiers, dogfood under load, residual risks, and one of five outcomes: accept, provisional, reject, subsume, or no new Primitive justified.

4.3 KFD-6: a bounded experimental loop

At loop step n, let E_n be an immutable causal-experience cut, O_n the declared ontology, G_n plural generation experiments, C_n a shared-budget method comparison, Q_n KFD-5 qualification, H_n held-out or independent evaluation, and A_n promotion authority. Draft KFD-6 proposes:

$$(E_n, O_n) \to G_n \to C_n \to Q_n \to H_n \to A_n \to O_n \text{ or } O_{n+1}. \tag{5}$$

The conservative baselines are part of the mechanism:

- **fixed ontology:** improve prediction, retrieval, or workflow while forbidding new first-class objects;
- **no new Primitive:** permit diagnosis and local abstraction but require the final decision to retain O_n; and
- **flattened history:** use anonymous events or endpoint facts under the same resource budget.

Candidates are evaluated for yield, false-candidate rate, ontology distance, compression, intervention value, transfer, falsifiability, cost, and auditability. Negative results remain part of the cut. Generated examples cannot be the only evidence. The generator cannot be the sole verifier or promotion authority.

4.4 The authority flow

The complete mechanism assigns a narrow authority to each layer:

Layer	Authority
Episode corpus	Evidence carrier with declared capture and stewardship; not an ontology decision.
KFD-4 replay	Declared projection and comparison; not source mutation or candidate certification.
Genesis method	Proposes a candidate and disconfirming test; not truth or promotion.
KFD-5 qualification	Judges bounded load-bearing utility against facts and alternatives; not universal adoption.
Held-out evaluator	Tests transfer and false candidates outside the discovery cut; not value or legal authority.
Promotion authority	Admits or rejects ontology change inside a declared scope; does not rewrite prior Episodes.

This separation is the central anti-self-certification property. It lets an agent contribute ontology change without turning model confidence into social or operational authority.

5 Problem-Definition Power and Knowledge Governance

Knowledge systems are often evaluated by answer quality while the space of permitted questions remains implicit. Yet a representation determines which objects can be counted, compared, assigned, challenged, or governed. The party that controls the durable schema can make some burdens visible and leave others as repeated human reconstruction.

Episode infrastructure changes this structure in two opposing ways. When it preserves multiple perspectives, causal links, responsibility, omissions, and portable evidence, later participants can

challenge the founding ontology. When it flattens histories into provider-controlled summaries or anonymous events, it can harden that ontology and make alternatives unverifiable. The ability to revise problem definitions is therefore not a magical property of memory. It is a governed capability distributed across capture, access, representation, replay, qualification, and promotion.

Boundary-object research explains how artifacts can retain identity across heterogeneous viewpoints without requiring consensus [15]. KFD adds an admission question: when does a recurring relation deserve responsibility as a Primitive, and who may decide? The answer cannot be "the model that named it." A participant has local epistemic priority over consequences it directly bears, but its wider interpretation remains a claim. Operators may steward shared infrastructure, but stewardship does not authorize appropriation of private experience or unilateral ontology change.

5.1 Rights and obligations

An Episode asset policy should make at least the following independently addressable:

- the right to know whether an Episode exists and which purpose justifies its retention;
- the right to inspect, export, correct through successor evidence, or request deletion where applicable;
- the right to see redaction, degradation, provider migration, derived indexes, and model-generated projections;
- the right of a consequence-bearing participant to contest a replay or candidate attributed to its perspective;
- the obligation to preserve prior cuts rather than rewriting history to match a later ontology; and
- the separation of access to evidence, authority to qualify, and authority to promote.

These are design obligations, not a claim that one policy fits every legal or cultural domain. Some Episodes should expire. Some should never be captured. Some may support only local replay, not cross-organization training or discovery. An honest system preserves these limits as part of the asset rather than treating them as obstacles external to it.

5.2 Why the issue is structurally important

If agents participate in producing candidate concepts, ontology change becomes part of the knowledge-production pipeline rather than a rare upstream design event. The pipeline can redistribute who notices unnamed burdens and who can propose new objects. It can also concentrate power in the corpus owner, model provider, or promotion authority. That is why the problem cannot be reduced to better retrieval. The object under design is a constitutional separation among history, interpretation, qualification, and admission.

The appropriate conclusion is neither "automate problem definition" nor "reserve concepts for humans." It is to let humans and agents contribute at any stage when their evidence and perspective are declared, while preserving independent challenge and accountable promotion for consequential changes.

6 Bounded Evidence from Current KFD Cases

6.1 Immutable evidence cut

This paper freezes KFD repository commit `ada402f671a878ff21e2f0d1c1b21c0458bac7a5`, dated July 21, 2026 [10]. Later changes cannot be projected backward into the paper's claim. At this cut, KFD-4 and KFD-5 are active; KFD-6 is draft. The live-case registry contains three active cases and six candidate tracks [9].

Live case	Track	Status	Claim boundary
Proof-Carrying Work Object	Pursuit	Provisional	Durable intended-change identity; no universal work ontology or accepted Primitive claim.
	Warrant	Provisional	Purpose-bound authorization; no replacement for enforcement or accepted Primitive claim.
Software-Work Perspective Settlement	Initiative	Accepted	Accepted only as a software-development Domain Profile Primitive.
	Assignment	Accepted	Accepted only as bounded software-work responsibility in the current profile.
	Project Cut	Accepted	Accepted only as a software-project settlement Primitive; not a fourth fact engine or universal object.
Decision-Admission Settlement	Consequential Settlement	Provisional	Draft KFD-11 procedure; no activation, universal minimality, or independent cross-domain adoption.

Table 2: KFD live-case evidence at commit `ada402f`. "Accepted" is scope-bounded and does not mean universal or historically novel.

The table establishes public case coordinates and outcome boundaries. It does not establish that an Episode corpus caused every candidate, that one method was necessary or sufficient, or that acceptance predicts transfer.

6.2 Two perspective transformations

The software-work case records two transformations over retained development history. Moving from Primitive developer to real human user exposed *Initiative* and *Assignment*: lower action coordinates did not by themselves state which work belonged together or whether a participant accepted responsibility. Moving from coordinator to successor acting agent exposed *Project Cut*: available source, Episode, Atlas, policy, omission, and risk records did not by themselves identify the one combination officially settled for continuation.

Each accepted track includes a deletion witness. If Initiative grouping is erased, equal lower objects can still require different coordination. If Assignment acceptance is erased, equal objective and

authority can still require different accountability. If Project Cut selection is erased, equal components can still distinguish a verified settlement from a candidate combination. The qualifications remain limited to their software profiles.

6.3 Natural pre-enactment of part of KFD-6

The Consequential Settlement case is the strongest evidence for this paper's specific thesis [5]. During real development, agents generated Claim, purpose-bound Assessment, and Decision before the maintainer became aware of those named objects. Later agents replayed retained development lineage, connected the intermediate ontology to Fact Admission, and compressed it into a provisional procedure. Maintainer authorization, repository review, and merge admission remained separate from generation.

This supports a bounded statement: an agent-generated ontology change can emerge under real work pressure, remain recoverable in retained history, and later receive separate KFD-5 qualification with less direct human translation than the founding pattern assumed.

It is not a conforming KFD-6 experiment. The experience cut was reconstructed after action. Methods and budgets were not preregistered. There was no plural method comparison against fixed-ontology and no-new-Primitive baselines. The evaluation was not held out from discovery, systematic false-candidate rejection was absent, and no autonomous trigger began ontology review. The case is a retrospective fixture for designing an experiment, not activation evidence.

6.4 Evidence grades

We use four grades to prevent rhetorical promotion:

E0 – repository fact.
Exact commit, schema, case, cut digest, status, test, review, or release coordinate.

E1 – implementation-backed case inference.
A bounded interpretation supported by retained public history and explicit falsifiers.

E2 – prospective hypothesis.
A mechanism or metric requiring a new controlled or shadow evaluation.

E3 – non-claim.
Universal validity, historical novelty, general autonomous discovery, or quantified social value not supported here.

Table 2 is E0. The natural pre-enactment interpretation is E1. The discovery option model and prospective evaluation are E2. A statement that KFD-6 is operationally proven would be E3 and is rejected.

7 Prospective Evaluation Protocol

A useful evaluation must distinguish three propositions: the Episode representation preserves information unavailable in flattened histories; the discovery mechanism uses that information better than conservative baselines; and the resulting candidate transfers without exporting more complexity or governance harm than it compresses.

7.1 Protocol A: representation ablation

Construct several immutable corpora from the same completed work Episodes:

1. full Episode assets with perspective, causality, responsibility, evidence, gaps, and stewardship metadata;
2. provenance-preserving events without participant consequence and ontology declarations;
3. anonymous event sequences with timestamps and payloads; and
4. endpoint Fact cuts only.

Blind evaluators to the known candidate names. Ask them to reconstruct why specified decisions differed, identify repeated unnamed burdens, and propose or reject candidate distinctions. Measure reconstruction time, unresolved questions, evidence alignment, candidate yield, false-candidate rate, and inter-evaluator agreement. If the full Episode condition does not improve recoverability or reduces it by imposing excessive structure, RQ1 is weakened.

The ablation must hold source cases and resource budgets constant. Otherwise a larger corpus or more capable model could be mistaken for a representation effect.

7.2 Protocol B: preregistered shadow KFD-6 loop

Before new work begins, freeze E_n and O_n, preregister method procedures, budgets, metrics, falsifiers, and autonomy boundaries, and reserve a held-out cut. Run at least two method families, for example perspective transformation and anomaly/reconstruction analysis, plus fixed-ontology and no-new-Primitive baselines.

Candidate names and attractive narratives are not the primary endpoint. The endpoint is whether a candidate survives KFD-5 minimum closure, alternative, deletion, fuse, falsifier, and dogfood gates and then improves a held-out decision. Retain rejected candidates and score why they failed. A system that generates many plausible objects but cannot reject them fails the experiment.

Shadow mode prohibits the loop from changing production ontology or participant responsibility. An independent evaluator receives the held-out evidence. A separate accountable authority decides whether any result warrants a bounded intervention.

7.3 Protocol C: transfer and intervention

A provisional candidate that passes Protocol B is mapped into at least one context outside its founding workstream. The test permits different names and representations. It asks whether an equivalent distinction:

- reduces repeated reconstruction or coordination cost;
- changes a decision that remained ambiguous under the prior ontology;
- preserves source Fact, Episode, perspective, and authority boundaries;
- admits a smaller rival representation if one works better; and
- remains removable or reversible when consequences disagree with the qualification.

Transfer within one organization's repositories is weaker than independent adoption. Both are useful but must not share a label.

7.4 Protocol D: governance and stewardship

The experiment audits the corpus itself. For sampled Episodes, verify integrity roots, retention basis, consent or other legitimate capture boundary, redaction, deletion behavior, export/import fidelity, provider migration, and the distinction between authoritative records and derived projections. Invite consequence-bearing participants to challenge replays attributed to them.

Governance failure can invalidate a technically promising result. A candidate is not qualified if producing it requires undeclared surveillance, prevents applicable deletion, silently broadens purpose, or makes independent challenge impossible.

7.5 Decision rule

The strong hypothesis survives only if the Episode condition produces a candidate that passes KFD-5, improves at least one held-out decision over both conservative baselines, rejects attractive false candidates, transfers across a declared boundary, and preserves governance constraints at acceptable cost. Failure should narrow the representation, mechanism, or domain claim. It must not be repaired by relabeling retrieval or clustering as Primitive discovery.

8 Related Work

Provenance and event history. W3C PROV provides a portable model for entities, activities, agents, derivation, and responsibility [12]. Process mining discovers and checks process models from event data and has long emphasized event-log quality and fitness between observed and modeled behavior [17]. KFD Episodes can use both. The additional problem is that the case, activity, responsibility, or process ontology may be the object under revision. KFD therefore requires the current ontology, observer, consequence boundary, and promotion path to remain explicit.

Causal discovery and representation learning. Causal inference distinguishes observation, intervention, and counterfactual reasoning [13]. Causal representation learning asks how high-level causal variables can be recovered from low-level observations and how they support transfer [14]. This is close to Primitive genesis, but KFD addresses a wider socio-technical object: a candidate may carry responsibility, authority, lifecycle, and participant consequences that cannot be promoted by model fit alone. KFD-5 qualification and separated promotion authority are therefore complementary to statistical causal evidence.

Episodic control and open-ended learning. Model-free episodic control stores high-value experience for rapid reuse [1]. Open-ended learning changes task distributions and objectives so agents continue acquiring behavior [16, 4]. These lines demonstrate the value of experience beyond one update and the difficulty of measuring continuing novelty. The present proposal is narrower in capability and broader in governance. It does not seek unbounded behavioral novelty; it tests whether a declared causal corpus supports a bounded ontology change that remains falsifiable and separately admitted.

Organizational memory. Organizational-memory research distinguishes acquisition, retention, retrieval, use, misuse, and abuse [18]. Episode asset stewardship shares this concern but rejects the repository metaphor when it implies a neutral past. Perspective and schema determine what enters memory, while later replay constructs a view under declared loss. The asset is thus a versioned evidence option, not a complete organizational recollection.

Boundary objects and distributed cognition. Boundary objects are adaptable to different social worlds while maintaining enough common identity for cooperation [15]. KFD's accepted software-domain candidates often mediate exactly such contact surfaces. The new contribution is a promotion discipline: adaptability and use do not alone establish a Primitive. The candidate must preserve a decision-changing distinction, survive alternatives and falsifiers, and keep its lower authorities inspectable.

Across these literatures, the common insight is that history, representation, and viewpoint shape action. The gap addressed here is an end-to-end contract for moving from governed causal history to a proposed ontology revision without merging generation, evidence, evaluation, and admission.

9 Limitations, Failure Modes, and Non-Claims

9.1 Evidence is first-party and retrospective

The current live cases are maintained by the same ecosystem that developed KFD. Public commits, immutable cuts, review records, and explicit falsifiers make the claims inspectable but do not remove correlated incentives. The Consequential Settlement mapping was reconstructed after action. Independent adopters and prospective trials are required.

9.2 Survivorship and narrative bias

Once a candidate is named, many histories can be retold as if they led inevitably to it. The mechanism counters this with frozen cuts, method plurality, conservative baselines, false-candidate retention, and held-out evaluation, but the current evidence does not yet demonstrate that those gates are sufficient.

9.3 Capture can distort action

Rich Episode capture may change participant behavior, create surveillance pressure, and privilege what is machine-recordable. A causal graph can still exclude consequences that occur outside the capture boundary. Declared gaps do not eliminate this harm. In sensitive settings, less retention or no capture may be the correct choice.

9.4 Ontology growth can export complexity

A false Primitive replaces local reconstruction with global schema, governance, migration, and training cost. Deletion and fuse tests are necessary, not sufficient. Progressive disclosure and rival representations must remain available. "No new Primitive" is a successful outcome.

9.5 No automatic epistemic justice

Perspective metadata does not guarantee that marginalized participants can contribute, contest, or control promotion. The corpus owner may still dominate access and interpretation. The governance requirements in Section 5 are research and design obligations, not a proof that a deployed institution is legitimate.

9.6 Explicit non-claims

This alpha paper does not claim:

- that KFD-6 is active or that a conforming autonomous loop exists;
- that every trace should become an Episode or every Episode be retained;
- that more data produces more Primitives;
- that perspective transformation dominates other genesis methods;
- that the six current candidate tracks are historically novel, universally minimal, or cross-domain requirements;
- that a schema validator proves causal truth, discovery, or promotion;
- that one physical object model is required; or
- that technical stewardship resolves legal, ethical, labor, cultural, or political authority.

These boundaries are part of the result. A theory of ontology revision that cannot preserve the possibility that it is wrong would reproduce the failure it claims to solve.

10 Conclusion

Historical assets are usually valued for remembering what a system already knew how to represent. This paper proposes a stronger and more demanding role: a governed causal Episode corpus can preserve the option to discover that the old representation itself was inadequate.

That option requires structure and restraint. KFD-4 makes replay perspective-bound and loss-aware. KFD-5 keeps candidate genesis open while making qualification answerable to facts, alternatives, deletion, falsifiers, and real work. Draft KFD-6 adds plural method comparison, fixed-ontology and no-new-Primitive baselines, held-out evaluation, false-candidate retention, and separated promotion authority. Episode storage alone performs none of these acts.

The current KFD case registry makes the research direction more visible. It contains six real candidate tracks, including three bounded software-domain acceptances and a retained natural pre-enactment of agent-generated ontology change. It also makes the missing evidence visible: no preregistered plural method experiment, no systematic false-candidate result, no held-out transfer, and no autonomous ontology-review trigger. That is why the evidence supports feasibility, not completion.

The practical consequence is significant. If agents can participate in defining the objects through which work is understood, problem definition becomes an operational knowledge-production surface. Preserving that surface can expand who notices unnamed burdens; governing it poorly can concentrate power in corpus, model, or promotion owners. The right response is neither unbounded retention nor unaccountable automation. It is a verifiable separation among causal history, perspective, candidate generation, qualification, independent evaluation, and admission.

The next milestone is therefore empirical: run the preregistered ablation and shadow-loop protocols in Section 7, retain the negative results, and let failure narrow the claim. If a candidate survives those gates and improves held-out continuation without hiding participant or source authority, Episode history will have demonstrated value not merely as memory, but as infrastructure for revising what a system is capable of knowing.

References

[1] Charles Blundell, Benigno Uria, Alexander Pritzel, Yazhe Li, Avraham Ruderman, Joel Z. Leibo, Jack Rae, Daan Wierstra, and Demis Hassabis. Model-free episodic control, 2016.

[2] Keren Dong. Facts, trust, and boundary primitives: A foundation model for real-world agent work. `https://github.com/kungfu-systems/paper-kfd-foundation-real-world-agent-work`, 2026. Alpha research artifact.

[3] Keren Dong. The observer is part of the system: Declared timelines for reliable human-agent work. `https://github.com/kungfu-systems/paper-observer-declared-timelines`, 2026. Alpha research artifact.

[4] Nicholas Guttenberg, Nathaniel Virgo, and Alexandra Penn. On the potential for open-endedness in neural networks. *Artificial Life*, 25(2):145–167, 2019.

[5] Kungfu Systems. Consequential settlement: Early kfd-6 feasibility case. `https://github.com/kungfu-systems/kfd/tree/ada402f671a878ff21e2f0d1c1b21c0458bac7a5/cases/live/decision-admission-settlement`, 2026.

[6] Kungfu Systems. Kfd-4: Views must remain bound to declared perspectives. `https://github.com/kungfu-systems/kfd/blob/ada402f671a878ff21e2f0d1c1b21c0458bac7a5/decisions/KFD-4.md`, 2026.

[7] Kungfu Systems. Kfd-5: Primitive discovery must separate genesis from qualification. `https://github.com/kungfu-systems/kfd/blob/ada402f671a878ff21e2f0d1c1b21c0458bac7a5/decisions/KFD-5.md`, 2026.

[8] Kungfu Systems. Kfd-6: Autonomous discovery must remain grounded in causal experience. `https://github.com/kungfu-systems/kfd/blob/ada402f671a878ff21e2f0d1c1b21c0458bac7a5/decisions/KFD-6.md`, 2026. Draft procedure.

[9] Kungfu Systems. Kfd live primitive-discovery case registry. `https://github.com/kungfu-systems/kfd/blob/ada402f671a878ff21e2f0d1c1b21c0458bac7a5/cases/registry.json`, 2026. Schema version 2 at commit ada402f.

[10] Kungfu Systems. Kung fu decisions: Causal experience and primitive discovery evidence cut. `https://github.com/kungfu-systems/kfd/tree/ada402f671a878ff21e2f0d1c1b21c0458bac7a5`, 2026. Commit ada402f671a878ff21e2f0d1c1b21c0458bac7a5, accessed July 21, 2026.

[11] Leslie Lamport. Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7):558–565, 1978.

[12] Timothy Lebo, Satya Sahoo, Deborah McGuinness, Khalid Belhajjame, James Cheney, David Corsar, Daniel Garijo, Stian Soiland-Reyes, Stephan Zednik, and Jun Zhao. Prov-o: The prov ontology. W3c recommendation, World Wide Web Consortium, 2013.

[13] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, second edition, 2009.

[14] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

[15] Susan Leigh Star and James R. Griesemer. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907–39. *Social Studies of Science*, 19(3):387–420, 1989.

[16] Adam Stooke et al. Open-ended learning leads to generally capable agents, 2021.

[17] Wil van der Aalst et al. Process mining manifesto. In *Business Process Management Workshops*, pages 169–194. Springer, 2012.

[18] James P. Walsh and Gerardo Rivera Ungson. Organizational memory. *Academy of Management Review*, 16(1):57–91, 1991.