

The Knowability Principle

A Civilizational Case for the Preservation of Consequential Decision Evidence

Author	Andrew D. Hackett
Version	1.0
Status	Foundational Proposal
Constitutional Status	Active
Supersedes	None
Superseded by	None
Ratification Process	Founder Ruling per HMOS-INDEPENDENT-VERIFIER-001
License	To be specified – proposed: open, permissive attribution

A future that cannot explain itself cannot improve itself.

1 EXECUTIVE SUMMARY

DECLARED ABSENT. *This paper operates at civilizational altitude and is itself of executive-summary length. No further compression is possible without loss of the substrate the paper is intended to establish. In accordance with the no-omission rule of the constitutional publication protocol, this section is declared absent with reason rather than silently omitted.*

2 ABSTRACT

Modern civilization has developed extraordinary capacity to generate, store, and act upon information. Yet across domains — medical, financial, scientific, legal, industrial, governmental — organizations continue to incur substantial cost reconstructing the evidentiary basis of consequential decisions after the fact. This paper proposes that this recurring cost reflects a first-order property of consequential systems that has not been named as such. It names the property *knowability*: the degree to which a consequential decision can be independently reconstructed by an observer who was not present at the decision, without recourse to institutional memory, proprietary infrastructure, or retrospective interpretation. The paper proposes The Knowability Principle:

*As consequential systems become more capable, the evidence they preserve
must become more durable than the memory they rely upon.*

The paper argues that this principle is a civilizational requirement rather than a product feature, and that the discipline it implies extends beyond software architecture into the general infrastructure of accountability. It is offered as the foundational proposition of the Decision Evidence Infrastructure Series.

3 INHERITANCE

None. This paper is the load-bearing root of the Decision Evidence Infrastructure Series. No prior paper in the series is required to read it.

4 RELATED WORK

Every civilization that has produced consequential decisions has eventually developed institutions whose purpose is the preservation of independently checkable evidence. Courts, notaries, coroners, regulatory commissions, aviation flight recorders, scientific peer review, historical archives, and public inquiry each emerged from a recognition that the underlying society could not operate a system whose past was unknowable. The intellectual tradition that has examined this pattern is broad and precedes computing.

This paper inherits from several distinct literatures without claiming to synthesize or replace any of them.

The philosophy of science — particularly Popper on falsifiability, Latour on inscription and the material substrate of scientific claim, and Hacking on the historical construction of evidence — establishes that evidence is not a property of events but of the artifacts that outlast them.

The law of evidence, from Wigmore forward and continuing into contemporary evidence scholarship, establishes that admissibility, weight, and independent verification are distinct properties, and that no evidentiary system can operate without articulating each.

The aviation and medical safety literatures — most notably the flight-recorder tradition and the adverse-event reporting frameworks initiated by Kohn, Corrigan, and Donaldson (1999) — establish that preserving evidence at the moment of consequential action reduces the cost of subsequent understanding, and that this reduction is measurable.

The cybernetics tradition, following Wiener and Ashby, establishes that observation and control cannot be cleanly separated, and that the substrate on which observation operates constrains the reasoning any observer can perform.

The provenance research literature in computing, including the W3C PROV standard and the tradition following Moreau, has developed formal representations of the evidentiary state surrounding data and computation.

The reproducibility literature in science establishes that reconstruction from published artifacts is often less complete than the community assumed, and that this incompleteness has measurable consequences for scientific progress.

The digital forensics literature (Casey and successors) establishes operational disciplines for evidence preservation in adversarial computing environments, though its assumptions do not generalize cleanly to the wider space of consequential systems.

This paper does not propose to replace any of these traditions. It proposes that the property they collectively address — *knowability* — deserves to be recognized as a first-order requirement of consequential systems in general, and that the operational discipline required to preserve it is currently unspecified as a distinct architectural concern.

5 DEFINITIONS

Consequential decision. A decision whose outcome materially influences subsequent states of a system, organization, or environment. Materiality is domain-specific and must be declared for any specific system.

Evidence. An artifact produced contemporaneously with a consequential decision that permits a subsequent independent observer to establish, within stated bounds, the state, inputs, authority, assumptions, and boundary conditions of that decision.

Reconstruction. The activity of establishing the evidentiary basis of a consequential decision after the decision has occurred. Reconstruction may draw upon evidence, upon institutional memory, or upon inference; the relative proportion determines the reconstruction's reliability.

Knowability. The degree to which a consequential decision can be independently reconstructed from evidence alone, without recourse to institutional memory, proprietary infrastructure, or retrospective interpretation.

Independent observer. An observer who was not present at the consequential decision, does not depend upon the systems that produced it, and reaches conclusions from the evidence rather than from testimony.

Bounded knowability. The recognition that no consequential decision is fully reconstructable. Every reconstruction has boundaries; the discipline of knowability requires those boundaries to be explicit rather than concealed.

6 OBSERVATIONS

This section reports empirical observations. It advances no conclusions.

Organizations across domains routinely conduct post-event reconstruction of consequential decisions. This is observably true of medical adverse events, financial irregularities, cyber incidents, industrial accidents, judicial proceedings, regulatory investigations, and public inquiries.

The cost of reconstruction is frequently substantial. In many publicly documented cases it exceeds the direct cost of the event being reconstructed. Public reports from major incidents describe reconstruction efforts measured in months of skilled human effort.

Reconstructions repeatedly depend upon evidence assembled from multiple systems, multiple owners, multiple formats, and multiple institutional contexts.

Reconstructions repeatedly encounter gaps that cannot be filled by any evidence currently available.

Reconstructions repeatedly rely upon institutional memory — the recollection of specific individuals — to fill gaps in preserved evidence. Institutional memory is empirically known to be selective, revisable, and lost when personnel change.

Contemporaneous preservation of evidence in specific domains (aviation, forensic accounting, clinical trials) is broadly associated with reduced reconstruction cost and increased independent verifiability in those domains, though the relationship is not universally quantified.

These observations do not establish why reconstruction is expensive or why evidence gaps persist. They establish that reconstruction, its cost, and its dependence on institutional memory are widespread across domains.

7 FIRST PRINCIPLES

This paper proceeds from three explicit first principles.

First Principle 1. *Every consequential decision occurs within an evidentiary state. At the moment a consequential decision is made, there exists a set of observations, authorities, constraints, assumptions, and uncertainties that jointly bear on the decision. This set may or may not be preserved. Its existence is prior to any question about its preservation.*

First Principle 2. *The evidentiary state of a consequential decision may be reconstructed by future observers. Reconstruction may serve operational, legal, safety, scientific, regulatory, historical, or moral purposes. The motive varies. The activity is constant across human institutions.*

First Principle 3. *Contemporaneous preservation of the evidentiary state produces a different reconstructive substrate than retrospective assembly. This principle does not claim that contemporaneous preservation is universally superior. It claims that the substrate produced by contemporaneous preservation differs in kind from the substrate produced by retrospective assembly. The difference in kind is what makes contemporaneous preservation an architectural concern rather than an operational preference.*

An earlier draft treated *contemporaneous preservation is generally more reliable than retrospective reconstruction* as a fourth principle. That statement is empirical rather than axiomatic and belongs in Section 6 as an observation rather than in this section. Its removal tightens the derivation in Section 8.

8 DERIVED CONSEQUENCES

The following consequences follow from the first principles.

Consequence 1. If reconstruction is a recurring activity across human institutions (Principle 2), and if contemporaneous preservation and retrospective assembly produce different reconstructive substrates (Principle 3), then the choice between them is architectural rather than incidental. It is a property of how the underlying system is built, not a matter of downstream diligence.

Consequence 2. The reconstructive substrate produced by contemporaneous preservation is qualitatively different from the substrate produced by retrospective assembly. It is bounded by what was actually preserved rather than by what can be inferred or recalled. Bounded reconstruction is auditable in a way that inferred reconstruction is not.

Consequence 3. If a society or system routinely produces consequential decisions whose evidentiary basis is not contemporaneously preserved, it accumulates a *reconstruction debt*. The debt is not visible until reconstruction is required, at which point it is paid at whatever cost is possible under the available substrate. Reconstruction debt shares the structural properties of other forms of debt: it compounds, it is invisible until called, and its cost is not paid by the party that incurred it.

Consequence 4. Absence is itself part of the evidentiary state. An observation that was expected but did not occur, an authority that should have been consulted but was not, an approval that never arrived — each of these carries evidentiary significance. A complete account of the evidentiary state must preserve both what was observed and what was expected-and-absent. The preservation of absence is not equivalent to the preservation of positive events, and it requires its own discipline.

Consequence 5. Independence of the evidence from the systems being examined is a structural requirement rather than a preference. Evidence whose verifiability depends on the honesty of the party being questioned cannot resolve disputes about that party's actions. This consequence is not new to computing; it is the reason courts do not accept a defendant's private notes as unimpeachable evidence.

Consequence 6. Where the substitution of narrative for evidence is permitted, it eventually occurs. Where the substitution of authority for accountability is permitted, it eventually occurs. Where the substitution of reputation for verification is permitted, it eventually occurs. These substitutions are not moral failures of specific institutions; they are structural consequences of operating without a substrate that resists them.

9 PREDICTIONS

If the framework advanced in this paper is correct, the following historical and empirical predictions follow.

Prediction 1. Domains that develop mature contemporaneous-preservation disciplines (aviation, clinical trials, forensic accounting, judicial evidence) exhibit lower reconstruction costs and higher independent-verification rates than domains where such disciplines are absent. The prediction is comparative across domains and is measurable in principle.

Prediction 2. As the capability and autonomy of consequential systems increases — particularly in artificial intelligence and autonomous physical systems — the cost of reconstruction rises faster than the cost of execution. Reconstruction becomes the binding constraint on accountability rather than execution capacity.

Prediction 3. Institutions that fail to develop contemporaneous evidence disciplines eventually substitute narrative for evidence, authority for accountability, and reputation for verification. This substitution is measurable and repeatable across historical cases.

Prediction 4. When contemporaneous evidence disciplines are introduced in previously undisciplined domains, reconstruction costs decline, independent verification rates rise, and post-event ambiguity decreases. These effects are measurable at the domain level over multi-year windows.

Prediction 5. The absence of a contemporaneous evidence discipline is more expensive than its presence in the long run, but the two costs fall on different parties and at different times. This displacement of cost is what makes the discipline structurally difficult to adopt in the absence of external pressure.

These predictions are intended to be evaluated empirically. They are stated in a form that admits measurement, at whatever precision the available data permit.

10 CONDITIONS FOR REFUTATION

The principle advanced in this paper would be materially weakened if independent evidence demonstrated any of the following.

That reconstruction costs across domains are not materially reduced by contemporaneous evidence preservation.

That domains without contemporaneous evidence disciplines exhibit accountability, independent verifiability, and reconstruction cost comparable to domains that have developed such disciplines.

That the historical substitution of narrative for evidence, authority for accountability, and reputation for verification does not follow from evidentiary absence but from independent causes not addressed by the discipline this paper proposes.

That the preservation of absence (Consequence 4) provides no reconstructive value beyond the preservation of positive events.

That evidence whose verifiability depends on the system being examined is empirically as reliable, in adversarial or contested cases, as evidence whose verifiability is independent.

That reconstruction debt (Consequence 3) does not accumulate under the conditions described, or accumulates at a rate materially lower than the paper implies.

These refutation conditions are stated explicitly because the principle is proposed for examination rather than for adoption on faith. Independent researchers are invited to design studies against these conditions.

11 KNOWN LIMITS

This paper does not claim any of the following.

That knowability is achievable in the absolute. Every reconstruction is bounded. The discipline requires those bounds to be declared, not eliminated.

That knowability guarantees correct action. A knowable system may still act badly. The principle addresses whether the action can be understood, not whether it was right.

That contemporaneous evidence preservation is universally cost-justified. The discipline is proposed for consequential decisions, and materiality is domain-specific. Not every decision is consequential; not every consequential decision merits equal preservation.

That the preservation of evidence is sufficient for accountability. Evidence is a substrate for accountability, not accountability itself. The paper proposes that evidence discipline is necessary; it does not claim it is sufficient.

That knowability replaces trust. It bounds the domains in which trust must be extended and the domains in which it may be verified.

That the principle is original to this paper. Each of its components is present in some form in prior traditions. What this paper proposes is the synthesis and the naming, not the invention.

12 FUTURE RESEARCH

The Knowability Principle names a first-order property of consequential systems. It does not specify the operational discipline required to preserve that property. Subsequent papers in this series propose specific architectural, metric, and infrastructural developments.

Paper FP-001, *The Cost of Knowing*, develops the economic and architectural case implied by this principle and introduces the *Time-to-Understand* metric as a candidate measure of knowability at the operational level.

Paper FP-002, *Decision Evidence*, fixes the vocabulary required to specify evidence preservation as a formal discipline.

Papers in Volume II develop the constitutional-systems layer, including the *Constitutional Substrate Layer* that governs how consequential reasoning systems load their evidentiary substrate before generating consequential outputs.

Volume III develops one specific implementation of the discipline.

Beyond the current series, the principle invites future research into: measurement methodology for knowability across domains; institutional patterns that historically produce or destroy knowability; the epistemic consequences of knowability discipline for scientific reproducibility, legal evidence, and public accountability; the philosophical relationship between preserved evidence and historical truth; and the extent to which knowability discipline can be applied to systems whose internal state is not directly observable.

13 CONCLUSION

The Knowability Principle proposes that the ability of a consequential system to be independently understood by observers who were not present at its consequential decisions is a first-order property of that system, and that the preservation of the evidence required for such understanding is a discipline distinct from the operational disciplines of execution, safety, security, and performance.

The principle does not name a product, a protocol, or a company. It names a governing requirement that any consequential system will eventually be measured against — whether by regulators, by courts, by successors, by history, or by the ordinary human necessity of understanding what happened.

The value of the principle will not be determined by the confidence of its author. It will be determined by whether the discipline it implies produces measurable improvements in reconstruction cost, independent verifiability, and reduction of ambiguity about consequential decisions.

As consequential systems become more capable, their evidence must become more durable than their memory.