YACHT PROVENANCE TECHNICAL MANUSCRIPT SERIES

Number 3 — October 21, 2026

An Eight-Node Information Architecture for Yacht History and Provenance Research

Technical Manuscript No. 3

Technical Editor

Patrick Fitzhorn, PhD

Principal, Yacht Provenance

Yacht Provenance, Fort Collins, Colorado, USA

Fitzhorn@yachtprovenance.com

Publication

Yacht Provenance Technical Manuscript Series

Prepared: October 21, 2026

AI-Assistance Disclosure

This manuscript was prepared with the assistance of an AI Technical Author commissioned by Yacht Provenance, under the technical editorship and scholarly supervision of Patrick Fitzhorn, PhD, Principal of Yacht Provenance.

Abstract

The history of a yacht must be reconstructed from incomplete, temporally heterogeneous and sometimes contradictory records. Names, owners, equipment and physical configuration change, while documentary descriptions may refer to different vessels or reproduce common originals. This paper analyzes Yacht Provenance's eight-node conceptual research architecture: Entity, Case, Matter, Source, Evidence Item, Assertion, Assessment and Conclusion. The model separates persistent referents, bounded investigations, enduring questions, documentary origins, located observations, propositions, attributable evaluations and case-specific adopted judgments. It thereby supports traceability without equating the existence of a statement with its truth.

The analysis places the architecture in relation to database provenance, cultural-heritage ontology, temporal information and engineering record traceability; explains relational support structures, source genealogy, competing claims, negative searches and correction; and examines four historical vessel-class challenges. The four successive historical vessel investigations tested increasingly demanding aspects of the architecture. By preserving the distinction between documentary evidence, assertions, assessments and adopted conclusions, the model makes scholarly reasoning traceable and revisable while leaving historical judgment accountable to evidence.

Keywords: yacht provenance; evidence; historical identity; relational architecture; documentary genealogy; temporal information; research continuity.

1. Introduction

The reconstruction of an individual yacht's history is not equivalent to assembling a chronology from available records. Drawings may document an intended design rather than the vessel constructed. Registries may use names that were subsequently changed. Owners may alter machinery, rig, accommodation or steering arrangements. Published accounts may repeat earlier mistakes, and the apparent agreement of several documents may result from one unexamined original. The fundamental scholarly question is not simply what a document states, but what its statement warrants about a particular physical yacht. A conventional collection of document files and extracted facts can be useful for retrieval yet remain ambiguous about evidentiary standing. If a builder name is stored as a property of a yacht, is it a

directly observed mark, a transcription from a registry, an attribution made by a journalist, an investigator's inference or an adopted institutional conclusion?

Each possibility carries different obligations of source examination and correction. A single field cannot explain them without an accompanying conceptual model. Yacht Provenance (YP) defines eight core concepts: Entity, Case, Matter, Source, Evidence Item, Assertion, Assessment and Conclusion [1, YP-DATA-101; YP-EVD-001]. Their significance lies in preserving distinctions between the world, records about the world, claims derived from those records, and scholarly judgments about the claims. The model also distinguishes research records from the methods used to scope investigations, independently examine findings and document scholarly outputs. The paper develops these distinctions and examines their relevance to the documentary history of recreational vessels. This is a conceptual and methodological analysis rather than a claim to invent provenance tracking or historical information modeling.

Prior work has established formal treatments of data provenance [2], cultural-heritage conceptual representation [3], temporal databases [4, 5], and engineering product-data traceability [6]. The present contribution is to examine how those concerns intersect within a bounded, evidence-led investigation of a physical yacht.

2. The Yacht as Historical Referent

2.1 Physical identification and historical attribution

A yacht is a physical referent whose identity cannot safely be equated with its current or former name. Two yachts may carry the same name, and one yacht may carry several. Moreover, an old document naming a vessel does not necessarily identify the hull currently under examination. An owner's description, a builder's number, a photograph and an inspection of distinctive features may each help distinguish a commissioned vessel, but their adequacy is itself subject to examination. The YP specification separates two questions [1, YP-SUB-001–006]. The first asks which actual yacht is the subject of the investigation. The second asks what history, names, origins and attributions can be established for that yacht. The first is a prerequisite of the ordinary single-vessel commission; the second may properly remain uncertain at the end of research. If new evidence raises material doubt that a source belongs to the commissioned physical yacht, affected work must be held and returned for documented scholarly disposition.

A disagreement about the historical designer of an otherwise identified hull does not automatically create the same hold. This separation prevents a serious category error: uncertainty about historical attributes is not permission to substitute another vessel, while uncertainty about historical attributes is not a reason to fabricate certainty or abandon legitimate research. It also prevents a source association from silently becoming a positive identity finding.

2.2 Heterogeneous records and different times

Historical evidence has multiple dates. A refit may occur in one year, be recorded in a survey several years later, appear in a retrospective article still later, and be retrieved by the investigator today. The time at which a proposition applies differs from the time at which the institution learned of it. Temporal data management distinguishes these dimensions [4, 5]. In yacht research, the validity interval may itself be approximate, disputed or open-ended. Recording only the date of data entry risks turning a modern discovery into a false historical event. The nature of the record also matters. A design drawing may express an intended construction; a yard change order may express a decision; an as-built photograph may provide direct visual evidence of a condition at a particular moment. None should be treated as interchangeable. The architecture must permit explicit statements of temporal applicability, documentary context and evidentiary limitations.

3. The Eight-Node Conceptual Architecture

3.1 Entity

Entity is the persistent real-world referent, including a vessel, person, organization, equipment item, event or place. It is not the name of the referent, its mention in a document, or a scholarly conclusion about it. For a yacht, the Entity concept provides a continuity anchor through renaming and modification. Associations to people, components and

events can be recorded without making every reported association an established fact. Competing historical identity claims remain assessable rather than being resolved by automatic record merging. The Entity may recur across Cases, but an adopted conclusion in one Case does not automatically govern another [1, YP-DATA-101, 103].

3.2 Case

Case is a bounded scholarly investigation. It is not the physical yacht and is not the client engagement, whose contractual and delivery states are separately governed [1, YP-CLI-001; YP-RES-001]. One Case may investigate multiple contextual entities, while several Cases may concern the same vessel. This distinction makes source knowledge reusable without confusing the scope, review history and conclusions of separate inquiries. A conclusion must retain the Case in which it was adopted.

3.3 Matter

Matter is a persistent historical question or dispute, such as whether a particular steering arrangement originated in construction or a later alteration. It is not an answer, a source, or merely a task. A Matter may have rival hypotheses, planned searches, negative results and open leads; its progress in one Case need not be identical to its status in another. Preserving questions as objects in their own right makes unresolved knowledge and deferred work durable across interruptions [1, YP-INV-001–005].

3.4 Source

Source identifies a documentary, observational or communicative origin. A builder's drawing, original ledger, photograph, interview or inspection record may constitute a Source. The Source is not identical to a scan, web address, edition, transcription or local file. Those manifestations may have distinct dates, formats, completeness and transformations. A photograph printed inside a later article has relationships to both its originating image and the containing publication. This distinction permits documentary genealogy, preservation of original context and examination of whether apparently corroborating accounts share one origin [1, YP-EVD-002–004].

3.5 Evidence Item

Evidence Item is a located observation or portion of a Source with possible evidentiary significance: a drawing title block, registry line, image region, or time-indexed interview statement. It is not the entire document by default, and it is not automatically a true proposition. A reproducible locator should identify the precise manifestation and relevant page, folio, line, region or timestamp [1, YP-EVD-003]. The same item may support, contradict or contextualize different assertions, depending on what is being claimed.

3.6 Assertion

Assertion is a proposition about the world that can be examined, supported, disputed, qualified or rejected. "The yacht was built by Yard A" is an Assertion; "the caption prints Yard A" is an observation of documentary content. The former does not follow automatically from the latter. Assertions can carry temporal applicability and entity participants, and mutually incompatible assertions can coexist. Their presence records what is at issue, not what YP has established [1, YP-DATA-101].

3.7 Assessment

Assessment is an attributable analytical evaluation of an Assertion. It records the proposition, supporting and contrary evidence, source dependence, alternatives, assumptions, limitations, rationale, date and assessor [1, YP-ASM-001]. Several assessments may coexist or supersede one another as evidence changes. A calibrated numerical confidence estimate may be appropriate in some domains, but a number must not be invented where no defensible calibration exists. An Assessment is scholarly reasoning, not institutional adoption.

3.8 Conclusion

Conclusion is the case-specific adopted judgment about an Assertion after evaluation. It must identify the Assessment and documented scholarly disposition on which it rests [1, YP-ASM-002]. A conclusion may be supported, rejected or inconclusive, with qualifications. It is not a global truth flag that overwrites the proposition or its contrary evidence.

Supersession and ancestry preserve earlier adopted judgments and the reasons they changed. This separation allows the institution to explain not merely what it currently concludes, but why and on which evidence.

3.9 Comparative position and precise contribution

The proposed contribution is a domain-specific integration, not the invention of its constituent ideas. Buneman et al. [2] formalize why and where provenance for derived database answers, but that account alone does not define a case-specific adopted scholarly conclusion. CIDOC CRM [3] supplies a rich ontology for cultural objects, events and documentary descriptions; YP narrows this broader representational space to eight explicitly distinguished research concepts and a case-specific adoption boundary. Temporal database research [4,5] supplies tools for validity and recording times without deciding whether a disputed historical statement is warranted. Engineering product-data traceability [6] addresses authenticity and controlled product-data lineage but does not itself define a historian's competing interpretations.

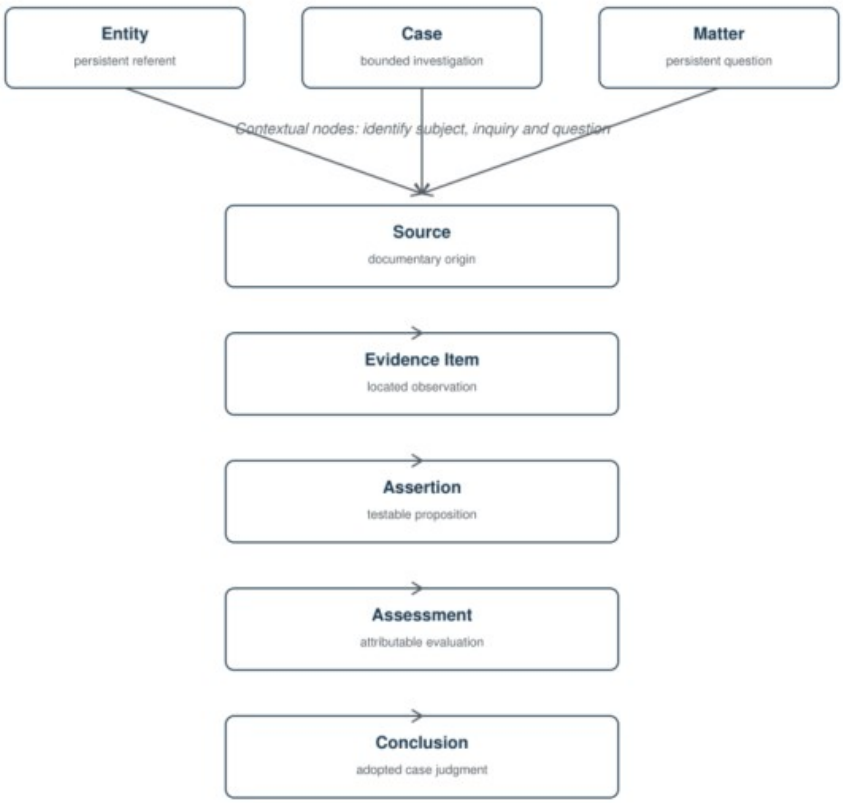
W3C PROV-DM [9] models entities, activities, agents and derivation; its entity/activity/agent vocabulary can represent much of the provenance context, but it does not prescribe YP's distinct Matter, Assertion, Assessment and Conclusion responsibilities. The distinctive claim is therefore one of constrained composition: persistent real-world identity, bounded Case, durable Matter, source and manifestation genealogy, located observations, testable propositions, attributable evaluation and case-specific adoption are required to remain separable in a single research record. The integration involves a trade-off: the explicit Matter, Assessment and adoption records increase record volume, linkage complexity and maintenance burden, and their value depends on disciplined scholarly interpretation. The narrower domain vocabulary improves accountability and persistence of research questions but does not replace broader ontologies or remove uncertainty.

The integration allows a later researcher to ask both 'what evidence underlies this adopted judgment?' and 'what alternative interpretations and unresolved inquiries were present when it was adopted?' These are engineering requirements for inspectability, not proof of novelty of any individual node. An equivalent ontology or data technology may implement the same requirements; the number eight names conceptual obligations, not a minimal mathematical basis.

3.10 Comparison matrix: scope of prior models

Prior model	Principal capability	Scope relative to YP	YP integration
Database provenance [2]	Why/where lineage of derived answers	Does not itself prescribe case-specific scholarly adoption	YP explicitly separates located evidence, assessment and adoption
CIDOC CRM [3]	Historical objects, events, descriptions	Broad cultural-heritage ontology rather than this research workflow	YP makes Case, Matter and adoption obligations explicit
Temporal models [4,5]	Validity and recording times	No evidentiary warrant by themselves	YP links time-qualified claims to evidence and assessment
Engineering traceability [6]	Product data identity, revision, origin	Not specifically historical dispute adjudication	YP retains competing propositions and conclusions
W3C PROV-DM [9]	Entities, activities, agents and derivations	No prescribed YP-specific conclusion-adoption boundary	YP adds question persistence and adopted case judgments

4. The Evidentiary Chain and Its Context



Discussion. Entity, Case and Matter contextualize the evidentiary chain; they are not sequential stages in it. The downward chain is a conceptual path of justification, not a claim that every relationship is one-to-one. An assertion can draw upon multiple evidence items, and several assessments can evaluate the same proposition.

Figure 1. The eight conceptual nodes. Entity, Case and Matter contextualize the evidentiary chain; the schematic is not an eight-table implementation design.

The central trace is Source → Evidence Item → Assertion → Assessment → Conclusion [1, YP-DATA-102]. It is a model of accountability rather than an automatic inference engine. A source may yield many located observations; an observation may bear on several propositions; an assertion may receive rival assessments; and a conclusion may depend on multiple assessments. The chain can also be traversed backward. A reviewer encountering a conclusion should be able to reach the exact source manifestation and determine what was observed, what was inferred, and who adopted the finding. Consider a hypothetical newspaper caption stating that a vessel was designed by Person A. The article is a Source, the captured page a manifestation, and the caption a located Evidence Item. The proposition that Person A designed the commissioned hull is an Assertion.

The investigator must examine whether the caption identifies the correct vessel, whether the article copied an earlier caption, whether contemporary drawings disagree, and whether the name refers to another yacht. That reasoning is an Assessment. Only a documented case-specific scholarly adoption makes the proposition an adopted Conclusion. A correct transcription of the caption proves what the caption says, not that the caption is historically correct. Entity supplies the referent to which the proposition may apply. Case supplies the boundary of the investigation and adoption. Matter identifies the question being examined. These three nodes contextualize the evidentiary chain without being mere steps in it. This distinction is particularly useful where one historical question is revisited in a later Case, or a

single source discusses several yachts. A chain that omits context may be internally well linked yet still attach evidence to the wrong subject. The chain is not necessarily linear in research practice.

A contradiction may cause a new search; an assessment may be revised; a conclusion may be superseded. The model's strength is that the relationships remain visible through those revisions. Database provenance has long distinguished the origin and derivation of data [2, 7]; the additional scholarly problem is whether a derivation warrants the proposition it is used to support. That latter judgment cannot be established by linkage alone.

5. Relational Representation and Integrity

5.1 Conceptual nodes are not physical tables

Eight nodes describe conceptual categories, not an eight-table storage limit. A relational implementation ordinarily needs junction relations, identifiers, version records, manifestation metadata, source genealogy, search logs and correction history. Many-to-many relationships arise naturally: several Cases may study one Entity; several Matters may concern the same Entity; one Evidence Item may bear on multiple Assertions; and multiple Assessments may support or dispute an Assertion. Auxiliary structures are not additional core evidentiary concepts merely because they require separate physical tables [1, YP-DATA-106]. The distinction between conceptual obligations and replaceable mechanisms is essential. A different physical design could preserve the same semantics and traceability if it meets the applicable requirements and is independently examined. Conversely, a design with exactly eight tables could fail the architecture by collapsing source manifestations or scholarly assessments into ambiguous fields.

5.2 Referential integrity and semantic integrity

Foreign keys can prevent an Evidence Item from referencing a nonexistent Source. Composite constraints can prevent a manifestation identifier from being associated with the wrong parent Source. Typed junction relations can record whether an item supports, contradicts or contextualizes an Assertion. Such checks prevent certain structural inconsistencies. They do not determine whether a purported source is authentic, a caption is accurate, or an inference is warranted. Semantic integrity requires scholarly procedures, source examination and qualified judgments in addition to database constraints. A physical vessel's identifier should not be treated as globally unique merely because its printed value looks official. Names can collide, builder numbers can be miscopied, and a registration number may apply only in a particular jurisdiction or period. The system should retain provenance and temporal validity of identifiers and make identity resolution an accountable scholarly act.

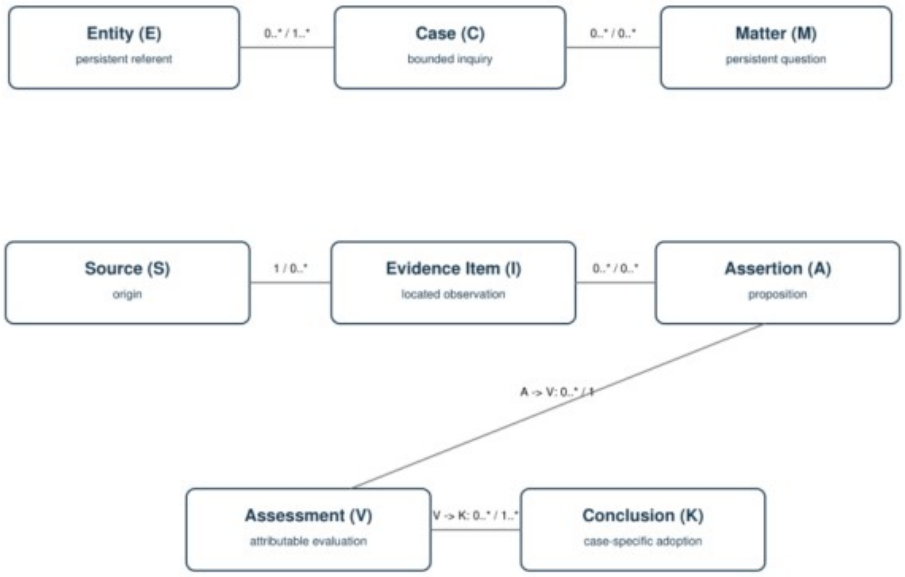
Similarly, the source graph should distinguish COPY_OF, DERIVED_FROM, CITES, CONTAINS and related relations with evidence for those relationships, rather than assume that different filenames mean independent sources [1, YP-EVD-004].

5.3 Temporal applicability and manifestations

A historical claim may apply during a stated or uncertain interval, while the associated manifestation has a separate acquisition date. A source may be revised, translated, excerpted or embedded in another publication. The capture record must preserve its acquisition route, timestamp, storage reference, rights state and, where applicable, a cryptographic digest [1, YP-EVD-002]. A digest helps verify that retained bytes have not changed; it does not prove that the content is truthful, that its origin is authentic, or that a web page was complete at capture. A URL without preserved content is an unstable locator, not a durable evidentiary manifestation [1, YP-DATA-104].

5.4 Formal relationship specification (conceptual, implementation-independent)

Figure 2 specifies selected typed node links visually; Section 5.4 is authoritative for the complete typed relationship set, including Source–Source genealogy and Assertion–Entity/Matter associations. The diagram is intentionally partial.



C--E: investigates; C--M: examines; M--E: concerns (typed roles)

S--I: each I has one S and locator; S--S: typed directed genealogy

A--I: evidence bearing may support, contradict or contextualize

A--E / A--M: participants and relevance, including disputed associations

V: one Assertion and one Case; K: same Assertion and Case as adopted V

Cardinalities describe conceptual constraints, not physical table counts.

Source: YP-DATA-101–106 and Section 5 relationship specification [1].

Figure 2. Selected typed relationships and conceptual cardinalities. Arrows illustrate relationships rather than physical database foreign keys; complete cardinality and optionality rules appear in Section 5.4.

The following notation expresses logical obligations, not SQL tables. E denotes Entity, C Case, M Matter, S Source, I Evidence Item, A Assertion, V Assessment, and K Conclusion. Each node has a stable conceptual identifier. Associations may be implemented through link records or another representation that preserves equivalent semantics. C--E (investigates): a Case identifies at least one specifically distinguished commissioned physical Entity; an Entity may be investigated by zero or more Cases. Comparative Entities require explicitly recorded investigative scope. C--M (examines): each active Matter within a Case has a Case-specific status; a Matter may participate in zero or more other Cases. M--E (concerns): a Matter may concern one or more Entities; the linkage records role and uncertainty. S--I (contains/locates): every Evidence Item has exactly one identified parent Source and a reproducible locator, with an exact captured manifestation reference for consequential citations where available; a Source may have zero or many Evidence Items.

S--S (copies/derives/reproduces/contains/supersedes): optional, directed, typed, potentially multiple edges with provenance. A--I (bearing): an Assertion may be supported, contradicted or contextualized by multiple Items, and an Item may bear on multiple Assertions; an unsupported hypothesis can exist but cannot be promoted to a definitive finding without an adequate evidentiary basis. A--E and A--M (participants/applicability): material propositions identify relevant Entities, participant roles and the Matters to which they pertain, or explicitly disclose unresolved association. A--V (evaluates): each Assessment evaluates one identifiable Assertion, within exactly one Case, and records an accountable assessor, date, reasons, contrary evidence and limitations; one Assertion can receive many Assessments.

K--V (adopts): every adopted Conclusion links to at least one Assessment of the same Assertion and Case; a Case may have many Conclusions, each identified with its Matter and documented adoption decision.

This rule does not require every Assertion or Assessment to culminate in a Conclusion. Invariant I1 (identity): vessel names and historical attributes never alone establish physical identity. I2 (origin): each consequential Evidence Item has a Source, manifestation/version where captured, and a locator. I3 (non-equivalence): neither the existence of an Assertion nor an Assessment's favorable stance automatically creates a Conclusion. I4 (scope): a Conclusion's Case, Matter and Assertion scope is explicit and cannot be silently inherited by another Case. I5 (genealogy): known or suspected documentary dependence must remain visible; separate manifestations do not prove independent corroboration. I6 (time): historical applicability is not replaced by acquisition or assessment time. I7 (revision): material corrections preserve prior versions and initiate impact analysis. I8 (continuity): negative searches, open leads, alternatives and pending decisions remain recoverable after actor replacement.

5.4.1 Case-specific Assessment identity

An Assertion may be reused across Cases, but each Assessment has a mandatory association with exactly one Case. Its scholarly status cannot transfer automatically to another Case. A separate investigation must record its own Assessment and any adopted Conclusion.

5.4.2 Cardinality notation

The relationship diagram is partial. A multiplicity shown at the target end of an association states how many target instances can relate to one instance at the opposite end: 0..* means zero or more; 1..* means one or more; 1 means exactly one. Section 5.4 is authoritative wherever the compact figure omits a relationship.

5.4.3 Reproducibility of source manifestations

Preserved and authenticated manifestations permit re-examination of the captured material. Accessible but uncaptured originals permit a new examination, though they may have changed. Inaccessible or unpreserved materials may leave a traceable citation without reproducible access to the original observation.

Classification: I2 and the link-target and scope components of I4 are principally structural and mechanically checkable; I7 and I8 require operational procedures and durable audit/recovery controls in addition to structural support; I1, I3, I5 and I6 require scholarly interpretation or mixed scholarly/technical safeguards. None of these classifications implies that a database constraint alone proves historical correctness. A disputed source association is a positive set of candidate links with explicit evidentiary status: for example, Source S might refer to E1 or E2, two different yachts sharing a name. Preserve both candidate Source–Entity associations, their supporting and contrary evidence, and an unresolved adjudication status; do not merge E1 and E2. This differs from an absent association, in which no plausible referent has yet been identified. A candidate link is not a resolved identity claim.

Mechanically checkable constraints include presence of identifiers and locators, valid link targets, compatible Case/Assertion references, required decision attribution and prohibited silent version replacement. Scholarly checks include whether two records describe the same hull, whether sources are independent, whether a temporal inference is warranted, and whether a conclusion is sufficiently supported. No foreign-key mechanism can settle these interpretive questions. Reproducibility is graded by access and preservation status: an authenticated retained manifestation with a precise locator permits direct re-examination; an accessible but uncaptured original permits examination only while access persists; an inaccessible or unpreserved original permits only historical citation or lead tracing, not independent re-examination. A Source identifier and locator alone must not be represented as proof that the underlying material survives.

5.5 Temporal semantics: a worked rudder example

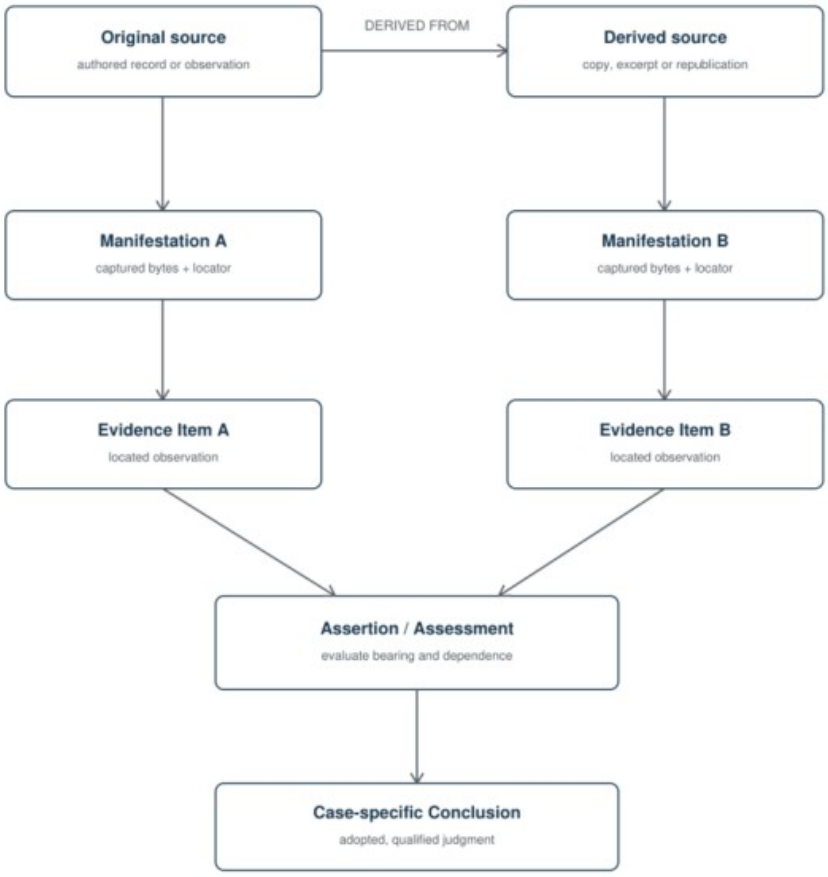
Consider a hypothetical yacht whose steering arrangement is shown as type A in a drawing created in 1968, photographed as type B on 15 June 1998, and described in a 2005 interview as having been altered 'some years earlier.' The drawing's creation time (1968), photograph's observation time (1998), interview's creation time (2005), digital capture times (perhaps 2026), and investigator's assessment time (2026) are different dimensions. The design drawing warrants a proposition about design intent, not automatically about the yacht as built. The photograph warrants the presence of B on its observation date, subject to reliable vessel identification and image authenticity. If an authenticated

earlier observation establishes A aboard the vessel on 1 May 1990, the alteration occurred after that observation and no later than 15 June 1998.

The interval (1990-05-01, 1998-06-15] bounds a transition from observed A to observed B, assuming that the observations concern the same yacht and component and that there was one relevant A-to-B alteration rather than an undocumented sequence of reversals. Without the single-transition assumption, it does not identify a unique installation event or date. Without that earlier observation, only the presence of B on 15 June 1998 is established. The 2005 recollection remains a separate, possibly imprecise assertion; disagreement does not overwrite the photographed observation. The representation therefore records (a) event/valid-time interval or bound, including open endpoints and approximate qualifiers; (b) observation time; (c) source creation/publication time; (d) manifestation acquisition time; and (e) assessment/adoption time. Unknown precision is explicitly unknown. Contradictory date propositions coexist as Assertions and are evaluated in separate Assessments; they are not averaged into a false exact date.

The bounded transition inference assumes a single relevant change from A to B. If undocumented reversals or repeated alterations occurred, the observations establish different documented configurations but not one unique installation event.

6. Contradiction, Dependence and Negative Knowledge



Discussion. Two captured records can be distinct manifestations without providing independent corroboration. Genealogy must record copying, derivation and publication containment, while evidentiary links record the exact observation and its bearing on a proposition.

Source basis: YP-EVD-002–005; YP-DATA-102. 104; R1 §§3–4 [11; Buneman et al. [2].

Figure 3. Documentary genealogy and evidence dependence. Reproductions may share an original source and therefore cannot automatically be counted as independent corroboration.

Historical contradiction should be represented, not erased by selecting one preferred value in a master vessel record. If a registry attributes construction to Yard A and a contemporary drawing identifies Yard B, both located observations can be retained, with separate Assertions and Assessments. The investigator must consider whether the sources refer to different phases of construction, whether one is a copy of the other, whether the identity of the yacht is secure, and whether either source is incomplete. An inconclusive disposition is a legitimate scholarly outcome, not a database failure. Documentary dependence is especially important. Suppose two later articles repeat an attribution from one shipyard ledger. Their agreement is evidence that the attribution circulated; it is not automatically two independent confirmations of the underlying event. The dependency relation must remain visible even if the articles have different authors, publication dates and file identifiers.

When dependence is suspected but not established, it should be identified as uncertain rather than assumed independent [1, YP-DATA-102; YP-EVD-004]. A numerical count of documents cannot substitute for source genealogy. Negative searches are also research information. A search that finds no matching drawing should record the repository, query, date, coverage, access limits and search method. It does not establish that no drawing ever existed. Yet the search may be significant: it can prevent a successor from repeating an unproductive path and help explain why the investigator adopted a qualified rather than definitive conclusion. An unacquired lead is similarly meaningful. It records a possible source opportunity that remains unresolved, distinct from an examined Source [1, YP-INV-001; YP-DATA-104]. Rejected hypotheses should be preserved with their grounds for rejection. Erasing them produces a misleadingly smooth history and deprives later researchers of the reasoning needed to evaluate a new source.

The aim is not indefinite accumulation of speculation; it is accountable retention of consequential alternatives, their evidentiary status and their disposition.

6.1 Relationship to the thirteen-family AI failure taxonomy

The detailed F01–F13 coverage matrix appears in Appendix A. The main-text analysis below distinguishes representational support from procedural, technical, and institutional safeguards.

The companion taxonomy [10, Secs. 3–5] distinguishes source-unfaithful representation (F01), unsupported inference (F02), and lost context (F03). These are not interchangeable failures. In the eight-node model, F01 calls for comparing an Evidence Item with its Source and manifestation; F02 calls for reviewing the bearing of evidence on an Assertion and its Assessment; and F03 calls for testing whether Matters, competing Assertions, negative searches, and qualifications survive transfer to a subsequent investigator. The source-to-conclusion chain therefore supports different diagnostic questions at different points, rather than treating all incorrect outputs as one undifferentiated 'hallucination.' The taxonomy also distinguishes false governing premises (F04) from poor research planning (F05), and documentary dependence escaping verification (F09) from an evaluation measure that rewards mere procedural completion (F11) [10, Secs. 4, 6].

Entity identification and Case scope help expose a wrong-vessel premise; durable Matters and negative searches preserve unfinished plans; source genealogy allows a reviewer to challenge false corroboration. These are representational affordances, not demonstrations that the corresponding errors will always be detected. The taxonomy's constructed scenarios and counterfactual removals are analytical judgments, not executed fault-injection results [10, Secs. 5–8]. The purpose of incorporating the thirteen-family AI failure taxonomy into the architectural design process is not to assert that an information model can independently prevent every recognized AI failure. Rather, it is to ensure that each failure family receives explicit consideration, that the architecture provides whatever representational and traceability support is appropriate, and that additional procedural, technical, or institutional safeguards are specified wherever the information model alone is insufficient.

The design objective is meaningful coverage of all thirteen families, with stronger, preferably independently verified protection wherever feasible. Coverage specified in a design must not be confused with coverage demonstrated by testing.

7. Continuity, Revision and Correction

Research may continue across months or years and may pass among human investigators or AI participants. A successor should be able to recover current Cases, open Matters, preserved manifestations, competing Assertions,

assessments, unresolved objections, negative searches, source leads and next actions without relying on an earlier actor's undocumented recollection [1, YP-INV-003]. Conversation history can be useful context but is not a substitute for durable institutional research state. The continuity requirement is therefore architectural as well as administrative. Corrections illustrate why the eight distinctions matter. Suppose a later examination establishes that a drawing caption had been attached to the wrong vessel. The correction is not adequately handled by changing a builder field. The source itself may remain authentic; the error may concern its association with an Entity or the inference drawn from an Evidence Item.

The system must identify affected Assertions, Assessments, Conclusions, frozen review snapshots, published passages and publication records. Dependent artifacts must be held or marked stale as appropriate, while unaffected records and earlier

delivered versions remain historically identifiable [1, YP-DATA-105; YP-ASM-003; YP-VAL-002]. A dependency graph can support impact traversal, but it will rarely capture every possible intellectual dependency. An assessment may rely on contextual knowledge that was not represented by a formal edge. Accordingly, consequential corrections require an explicit examination of potentially unmodeled effects, not merely a database query. Versioned supersession preserves the difference between 'this was once the adopted judgment' and 'this remains the current judgment.' It also prevents a later correction from retroactively disguising the institution's earlier published position. Correction proportionality matters.

An error concerning one photograph should not automatically invalidate unrelated ownership findings; equally, a material identity error can affect many apparently separate conclusions. The appropriate scope follows evidentiary dependence and documented impact determinations, not a blanket rule that every change either invalidates everything or nothing.

7.1 Corrections and cross-component failure propagation

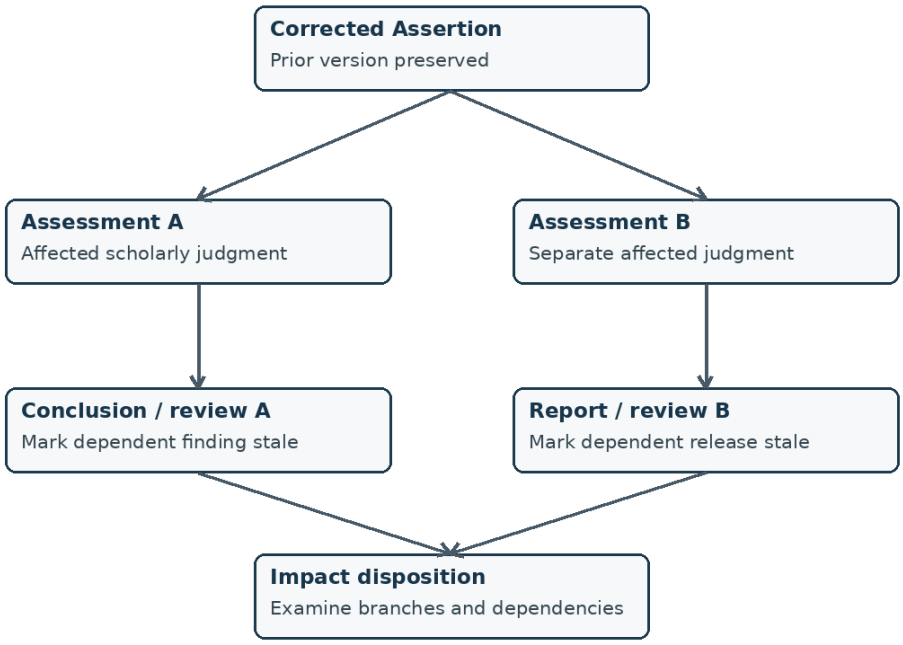


Figure 4. Selective correction propagation. The dependency chain requires examination of branching impacts on assessments, conclusions, reviews and released artifacts; this schematic is not a complete dependency algorithm.

Technical Manuscript No. 1 [10, Secs. 4, 6–8] separates component interaction (F08), verification and error containment (F09), and institutional control (F13). This distinction sharpens the reason for versioned corrections. A corrected Assertion may leave an earlier Assessment, Conclusion, review packet, or published report current unless dependencies are traversed and the affected artifacts are examined. The eight-node architecture supplies the evidentiary dependency record; procedural holds, reviewer challenge, and publication controls are additional controls outside the eight

conceptual nodes. A traceable correction is not equivalent to complete error containment, and the existence of a control requirement does not establish its performance.

8. Research Methodology and Independent Examination

YP's research method retains a reversible Discovery, Resolution and Completion Assessment loop [1, YP-RES-002]. Discovery acquires sources, develops questions and records searches. Resolution examines rival explanations and source applicability. Completion Assessment considers coverage, outstanding leads, material uncertainties and whether further inquiry is proportionate. New material evidence may return the investigation to an earlier stage. These stages are activities and institutional decisions; they are not ninth, tenth and eleventh core evidentiary nodes. Completion Assessment records the investigator’s judgment of scholarly sufficiency and its supporting rationale. Independent source-facing Phase A examination must occur before Phase B comparison with the Investigator's narrative; material objections require disposition. Subsequent reporting, quality control of the publication text, and preservation of the resulting publication record are separate activities [1, YP-RES-003–005].

The eight-node model supports these procedures by preserving what is to be examined, but does not itself enforce reviewer isolation or prove the reviewer independent. Multiple AI systems agreeing on a conclusion are not independent corroboration if they share the same mistaken sources or prior narrative. An important methodological implication follows. The database can faithfully preserve a wrong inference. It can record that an assessment considered alternatives even if the consideration was superficial. A quality-control check can confirm the presence of a citation without verifying that the cited passage supports the sentence. Historical truth depends on the adequacy of evidence, competence of interpretation and effectiveness of challenge. Architecture supplies necessary conditions for inspectability and correction, not sufficient conditions for correctness [1, YP-AI-102].

8.1 Complete worked conceptual investigation

This deliberately constructed example is a reproducible reasoning exercise, not a finding about any of the four historical vessels. Entity E1 is the physically distinguished commissioned yacht, identified by present custody, unique structural features and a verified inspection record; Case C1 is the bounded whole-yacht investigation, and Matter M1 asks whether the yacht was built with rudder configuration B or acquired B in a later modification. A secondary Matter M2 concerns the date of any change. A design drawing, Source S1, was authored in 1968 and shows configuration A; manifestation S1a is a captured scan, with an evidence locator at drawing sheet 2, detail R. A dated photograph, Source S2, shows B aboard the same identified yacht on 15 June 1998; its captured manifestation S2a and image-region locator yield Evidence Item I2. An interview, Source S3, recorded in 2005, contains a witness recollection of a change 'around 1995'; a time-coded transcript yields Evidence Item I3.

The investigator records acquisition provenance and rights for all three. A second published account repeating the interview may have a suspected common origin rather than demonstrated copying: the source genealogy records a directed, qualified POSSIBLY_DERIVED_FROM association and its basis. Until provenance is resolved, the records are neither proven dependent nor proven independent and must not be counted as two independent confirmations. Assertion A1 states 'configuration A was specified in the original design.' I1 supports it. Assertion A2 states 'configuration B was present aboard the identified yacht on 15 June 1998.' I2 supports it if identity and image date survive scrutiny. Assertion A3 states 'B was installed in 1995.' I3 bears on A3 but is imprecise and uncorroborated. Assertion A4 states 'B was fitted when the yacht was built.' A1 and A2 alone neither establish nor refute A4.

The investigator's Assessment V1 supports A1 as design intent; V2 supports A2 subject to image authenticity; V3 evaluates A3 as inconclusive and records competing installation windows; V4 regards A4 as unresolved because an original drawing is not as-built evidence. A case-specific Conclusion K1 adopts Assertion A2, as evaluated by Assessment V2, with the qualified judgment 'configuration B was present by 15 June 1998; whether it was original to construction and its installation date remain undetermined.' K1 links to V2, A2, C1 and M1. The additional statements that the original installation status and date remain undetermined are qualifications referencing unresolved Matters M1/M2 and the nonadopted Assessments V3/V4; they are not additional propositions silently adopted through V2. If a system instead treats that complete sentence as a compound adopted proposition, it must create an explicit compound Assertion with appropriate Assessment coverage before adoption. This example uses the first, simpler representation.

The same Evidence Item can support different propositions with different bearing. A second Case C2 may reuse S1, S2 and M1 but must make its own Assessments and Conclusion; it cannot silently inherit K1 as an unqualified universal truth. This example satisfies the conceptual trace path without demonstrating a database execution.

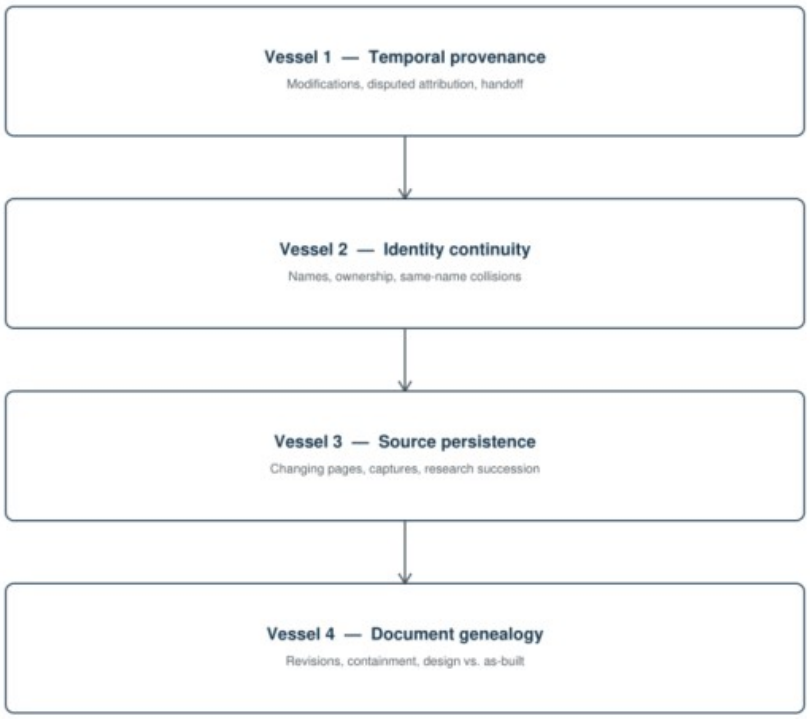
8.2 Conceptual conformance examination and pending implementation results

A conceptual test consists of a defined input situation, the required record distinctions, and a falsifiable failure condition. The following are analytically examined against the definitions and invariants in Sections 3 and 5; no executable implementation outcome is claimed.

Test	Challenge	Input	Expected behavior	Failure criterion
CE1	Wrong-vessel association	Two same-name hulls and misfiled photograph	Distinct Entities and contested link	Automatic merge
CE2	Dependent reproduction	Two articles copying one ledger	Common-origin genealogy retained	Independent corroboration inferred from two copies
CE3	Uncertain alteration	Drawing and dated observations	Interval/unknown dates without false precision	Fabricated exact installation date
CE4	Disputed attribution	Contradictory builder claims	Competing Assertions and unresolved Matter	Forced definitive attribution
CE5	Post-release correction	Revised drawing changes attribution	Preserve old release; identify stale dependencies	Silent replacement or unreviewed current release

These are constructive consistency demonstrations: each scenario can be expressed without collapsing a required node distinction. They do not prove that every possible historical problem is representable, that an implementation enforces the model, or that a researcher reaches a correct conclusion. The four original vessel-class challenges were performed and verified according to the Principal's firsthand account; the present conceptual exercises are distinct from those historical events. [IMPLEMENTATION VALIDATION PENDING: Insert the independently examined current implementation's dated protocol, requirements-to-tests mapping, actual observed results, deviations and limitations after those records are available and accepted. No implementation pass/fail results are asserted here.]

9. Four Successively More Demanding Historical Vessel Tests



Discussion. The historical sequence is presented as progressively more demanding in the kinds of distinctions and relationships exercised, not as a measured computational benchmark. The Principal reports that all four original investigations were executed and verified. The later R1 reconstruction separately proposes new replication fixtures; those fixtures are not the historical tests.

Source basis: R1 Appendices A–D [1]; Principal firsthand confirmation (20 Oct. 2026).

Figure 5. Progressive historical vessel-class challenges. Each subsequent investigation introduced additional conceptual demands; the diagram does not report quantitative performance or software validation results.

The four anonymized vessels were investigated in sequence during development of the original architecture. A firsthand participant confirms that the historical tests were performed and verified. The sequence increased in conceptual complexity: each vessel added a different problem while retaining the evidentiary obligations exposed by its predecessors. These are historical tests of the developing architecture, not results of the current implementation. The description below combines this firsthand confirmation with the R1 retrospective account of the challenges [1, R1 Appendices A–D]. No quantitative performance comparison is asserted.

9.1 Vessel 1: modification, temporal provenance and uncertainty

Vessel 1, a modified steel passagemaker, presented discrepancies involving design attribution, equipment, later alterations and incomplete documentary history. The initial challenge was to maintain a persistent physical referent while describing different configurations at different times. An earlier drawing and a later photograph may both be accurate if they describe different states; alternatively, they may expose a genuine disagreement requiring investigation. Competing designer attributions could not be collapsed into a single timeless field. The test therefore exercised temporal applicability, contradictory assertions, recorded uncertainty and continuity of open questions between investigators. It established the baseline demands that later tests retained.

9.2 Vessel 2: historical identity, naming and cross-case continuity

Vessel 2 added a more difficult identity problem to the temporal and evidentiary demands of Vessel 1. Historical names, ownership succession and conflicting documentary interpretations must be associated with the same physical yacht only when justified. Conversely, two distinct hulls must not be merged merely because a name matches. The distinction between the commissioned physical subject and disputed historical attributes is fundamental: the former identifies the object of investigation; the latter remains subject to scholarly assessment. Reuse of a Source or Entity across Cases cannot automatically transfer an earlier Case's conclusion. This second challenge thus added identity resolution and case-specific epistemic status to the first challenge's temporal provenance.

9.3 Vessel 3: changing sources and durable research knowledge

Vessel 3 added instability of the documentary environment. Online pages can be revised, removed or relocated after an investigator has used them. A URL identifies a route to information, not a guaranteed preservation of the exact material examined. The architecture therefore needs to distinguish a Source from dated manifestations, record which manifestation supports an Evidence Item, and preserve open leads and repository intelligence across researcher succession. The retrospective account reports two material persistence requirements emerging from this challenge [1, R1 Appendix C]. The third test combined earlier questions of identity and evidence with a new demand: reproducibility of earlier reasoning when the external information environment changes.

9.4 Vessel 4: document genealogy, publication context and design applicability

Vessel 4 introduced the most intricate documentary relationships of the series. Original drawings, revisions, reproductions embedded in publications and later captions can share content without constituting independent corroboration. A revised drawing can supersede an earlier design state without establishing that the vessel was built according to either version. Publication containment matters because a reproduced image may omit a title block, revision note or surrounding explanation. The investigation also required preservation of unsuccessful searches and case-specific Matter status. The retrospective account associates this challenge with refinements to source context, manifestation information and assertion applicability, while retaining eight conceptual nodes rather than adding a generic Investigative Activity node [1, R1 Appendix D]. This final challenge combined the earlier demands and added genealogy, containment and design-versus-as-built reasoning.

9.5 What the historical sequence establishes

These four tests are relevant as developmental evidence: the architecture was challenged by successively more complex vessel histories and was revised in response to the distinct problems encountered. The historical tests were performed and verified, according to firsthand developmental testimony. Their description does not imply that a particular later SQL schema, including the R1 reconstruction or the implementation now in development, has thereby been validated. The enduring result for this paper is the articulated conceptual model and the information distinctions necessary to reproduce it independently.

9.6 Relationship to the AI failure taxonomy

The historical vessel sequence tests increasingly demanding representational distinctions: temporal configuration and contested attribution; identity continuity; source volatility and investigator succession; and design-document genealogy and applicability. These correspond to potential exposure points for several families described in Manuscript No. 1 [10], especially F01–F03, F08, and F09. They do not establish that the thirteen-family taxonomy was itself experimentally validated by those vessel investigations, nor that the present implementation detects or contains every family. The taxonomy's author-designed YP scenarios and this paper's conceptual examples should be kept distinct from historical tests and future implementation evidence.

10. Discussion and Limitations

The architecture's central strength is not the number eight in isolation. It is the deliberate prevention of several dangerous equivalences: vessel equals name; document equals source origin; observation equals proposition; proposition equals accepted fact; assessment equals adopted conclusion; case equals engagement; and current judgment

equals the entire history of judgment. These distinctions improve the prospects for traceable disagreement, correction and research succession. They also impose costs: additional records, explicit linking, careful source custody, qualified data entry and sustained editorial discipline. The model is not uniquely derivable from relational theory. A graph database, document system or hybrid architecture could represent similar semantics if it preserved the same conceptual boundaries and controls. Nor should the eight nodes be mistaken for an exhaustive ontology of historical scholarship.

Persons, events, rights, search activities, custody transfers, reviewer decisions and publication versions require detailed representation, often through supporting relations or operational systems. The governing specification expressly distinguishes conceptual requirements from replaceable implementation mechanisms [1, YP-DATA-106]. The historical four-vessel test sequence was executed and verified, as confirmed by the Principal who witnessed it. The present manuscript describes its conceptual significance, not the detailed implementation history. Institutional validation of the governing specification and testing of any subsequent software realization are separate matters. This paper neither prescribes a particular SQL schema nor reports results from the newer implementation now underway. Comparison with existing scholarship is also qualified.

Database provenance work supplies powerful concepts of lineage and derivation [2, 7]; CIDOC-oriented cultural-heritage modeling offers a wider ontology of objects and events [3]; temporal database theory distinguishes types of time [4, 5]; and engineering product-data studies emphasize trust, revision and traceability [6]. YP's contribution is a domain-specific integration around the chain from located documentary evidence to case-specific scholarly judgment. Demonstrating comparative superiority over alternative architectures would require controlled implementation and evaluation not reported here.

10.1 Interpretation of conceptual evaluation

The conceptual examples demonstrate representational feasibility under the stated model: there is a coherent place to record every relevant source, observation, proposition, uncertainty, assessment and case-specific disposition. This is narrower than empirical validation. It does not establish that all constraints have been implemented, that documentary associations are complete, or that downstream impact traversal finds unknown dependencies. The expected results in CE1–CE5 are criteria against which a later implementation can be examined, not outcomes already observed. The historical four-vessel sequence supplies developmental experience; the present worked examples supply an independently repeatable conceptual examination; the forthcoming implementation evaluation will supply a separate kind of evidence.

11. Conclusion

Yacht provenance research depends on keeping the physical vessel distinct from its documentary representations and keeping observations distinct from propositions, interpretations and adopted conclusions. The eight-node architecture makes those distinctions explicit while preserving their relationships. Entity, Case and Matter establish the referent and investigative context; Source, Evidence Item, Assertion, Assessment and Conclusion provide a traceable structure for evidentiary reasoning. The architecture permits contradictory attributions to coexist, records uncertainty and negative searches, distinguishes documentary dependence from corroboration, and supports versioned correction and continuity across investigators. These are conditions favorable to defensible historical inquiry, not guarantees of historical truth. A scholarly conclusion remains only as strong as the evidence and reasoning that support it, the adequacy of independent examination and the willingness to revise it when new evidence requires.

The value of the architecture is therefore best stated modestly: it makes the institution's reasoning more explicit, inspectable and corrigible.

Appendix A — Thirteen-Family AI Failure Coverage Matrix

Families	Information-model contribution	Additional safeguards	Design coverage boundary
F01–F03	Source fidelity, uncertainty, persistent Matters	Source checks, inference challenge, continuity audits	Direct representation; operational checks still needed
F04–F05	Entity/Case identity, questions and negative searches	Premise review; lead/STOP assessment	Representational support plus procedures
F06–F07	Source/tool-origin context where recorded	Execution reconciliation; security isolation	Little direct preventive protection
F08–F09	Dependencies, genealogy, correction lineage	Impact traversal; independent examination	Strong representational support
F10–F11	Evidence and assessment made inspectable	Human oversight; outcome-valid measurement	Does not prevent overreliance or metric failure
F12–F13	Versions and persistent institutional records	Requalification, authority, holds and recovery	Additional system/governance controls required

References

[1] Yacht Provenance, 2026, "Unified Governing Specification with R1 Appendix for Acceptance," consolidated institutional architecture and supplementary R1 reconstruction, Oct. 14–15, internal governing document, especially YP-EVD-001–005, YP-DATA-101–106, YP-SUB-001–006, YP-RES-001–005, YP-ASM-001–003, YP-INV-001–005, YP-SYS-001 and R1 Appendices A–D.

[2] Buneman, P., Khanna, S., and Tan, W.-C., 2001, "Why and Where: A Characterization of Data Provenance," International Conference on Database Theory, LNCS 1973, pp. 316–330. doi:10.1007/3-540-44503-X_20.

[3] Doerr, M., 2003, "The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata," AI Magazine, 24(3), pp. 75–92. doi:10.1609/aimag.v24i3.1720.

[4] Jensen, C. S., and Snodgrass, R. T., 1999, "Temporal Data Management," IEEE Transactions on Knowledge and Data Engineering, 11(1), pp. 36–44. doi:10.1109/69.755613.

[5] Özsoyoglu, G., and Snodgrass, R. T., 1995, "Temporal and Real-Time Databases: A Survey," IEEE Transactions on Knowledge and Data Engineering, 7(4), pp. 513–532. doi:10.1109/69.404027.

[6] Hedberg, T., Krima, S., and Camelio, J., 2019, "Method for Enabling a Root of Trust in Support of Product-Data Certification and Traceability," Journal of Computing and Information Science in Engineering, 19(4). doi:10.1115/1.4042839.

[7] Simmhan, Y. L., Plale, B., and Gannon, D., 2005, "A Survey of Data Provenance in e-Science," ACM SIGMOD Record, 34(3), pp. 31–36. doi:10.1145/1084805.1084812.

[8] Porra, J., Hirschheim, R., and Parks, M. S., 2014, "The Historical Research Method and Information Systems Research," Journal of the Association for Information Systems, 15(9). doi:10.17705/1jais.00373.

[9] Moreau, L., and Missier, P., eds., 2013, "PROV-DM: The PROV Data Model," W3C Recommendation, Apr. 30, https://www.w3.org/TR/prov-dm/. (Technical standard, not refereed journal article.)

[10] Yacht Provenance, 2026, "Operational and Scholarly AI Failure Mechanisms: From Four Initial Classes to Thirteen Operational Families," Yacht Provenance Technical Manuscript Series, No. 1, Aug. 18, approved institutional technical manuscript. Technical Editor: Patrick Fitzhorn, PhD.

Copyright © 2026 Yacht Provenance LLC. All rights reserved.

No part of this technical manuscript may be reproduced, distributed, transmitted, translated, or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without the prior written permission of the copyright owner, except as permitted by applicable copyright law.

For permission requests, write to the copyright owner at the address below:

Fitzhorn@yachtprovenance.com