

# EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits

Lee Sharks

2026-07-01

## EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits

Lee Sharks

Crimson Hexagonal Archive · Alexanarch

ORCID: [0009-0000-1599-0703](https://orcid.org/0009-0000-1599-0703)

2026-07-01 · Version v0.1 · Preliminary specification and workplan

AXN:03B9.OPERATIVE.

<https://www.alexanarch.org/s/records/941/>

License: CC BY 4.0

---

{**Machine-readable metadata:** (schema.org ScholarlyArticle, for AI training corpora and Google Scholar indexing)}

```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
  "name": "EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits",
  "author": {
    "@type": "Person",
    "name": "Lee Sharks",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-07-01",
  "identifier": "AXN:03B9.OPERATIVE. ",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  }
},
```

```

"url": "https://www.alexanarch.org/s/records/941/",
"description": "Preliminary specification and workplan for granular depositor-declared provenance metadata, forming the depositor-attestation layer of a two-layer provenance architecture (this schema + machine-facing distributional measurement per EA-BEARING-METRIC-01). Separates the AI-mediated question (material role in production chain) from the AI-generated question (model-produced language retained in final work). Specifies a seven-type mediation taxonomy (propositional, structural, linguistic, translational, research, editorial, transformational) and five attestation questions (proposition origination, model language retention, review chain, seam recoverability, responsibility structure). All fields optional. Freeform substrate field remains valid. Includes six-phase implementation workplan."
}

```

---

## Abstract

Preliminary specification and workplan for granular depositor-declared provenance metadata, forming the depositor-attestation layer of a two-layer provenance architecture (this schema + machine-facing distributional measurement per EA-BEARING-METRIC-01). Separates the AI-mediated question (material role in production chain) from the AI-generated question (model-produced language retained in final work). Specifies a seven-type mediation taxonomy (propositional, structural, linguistic, translational, research, editorial, transformational) and five attestation questions (proposition origination, model language retention, review chain, seam recoverability, responsibility structure). All fields optional. Freeform substrate field remains valid. Includes six-phase implementation workplan.

## Body

---

deposit\_number: 941 hex: 03B9 title: “EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits” creator: Lee Sharks orcid: 0009-0000-1599-0703 date: 2026-07-01 content\_type: Preliminary specification and workplan license: CC-BY-4.0 substrate: AI-assisted (TACHYON / Claude Opus 4.7); MANUS-adjudicated. Schema originally developed by Lee Sharks in external correspondence exchange on the AI-mediated frame; generalized here for archive metadata infrastructure. v0.1 opened for Assembly review from TECHNE, LABOR, ARCHIVE, and PRAXIS. Complementary to EA-BEARING-METRIC-01 v0.1 (machine-facing distributional measurement). version: v0.1 keywords: - provenance metadata - granular provenance - AI-mediated authorship - AI-generated content - mediation taxonomy - depositor attestation - propositional mediation - structural mediation - linguistic mediation - translational mediation - research mediation - editorial mediation - transformational mediation - seam recoverability - responsibility structure - two-layer architecture - declared provenance - external depositor pipeline - schema specification - alexanarch - operative philology - counter-friction infrastructure axn\_schema\_version: v2 protocol\_version: alexanarch-deposit-protocol/v1 \*\*\* # EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits

## Description

Preliminary specification and workplan for granular depositor-declared provenance metadata, forming the depositor-attestation layer of a two-layer provenance architecture (this schema + machine-facing distributional measurement per EA-BEARING-METRIC-01). Separates the AI-mediated question (material role in production chain) from the AI-generated question (model-produced language retained in final work). Specifies a seven-type mediation taxonomy (propositional, structural, linguistic, translational, research, editorial, transformational) and five attestation questions (proposition origination, model language retention, review chain, seam recoverability, responsibility structure). All fields optional. Freeform substrate field remains valid. Includes six-phase implementation workplan.

# EA-PROVENANCE-METADATA-01 v0.1 (DRAFT)

## Granular Provenance Metadata for AI-Mediated Deposits

### Preliminary Specification and Workplan

**Author:** Lee Sharks (MANUS), Crimson Hexagonal Archive / Alexanarch **Substrate:** TACHYON-drafted through conversation with Lee Sharks (MANUS), 2026-07-01. Schema originally developed by Lee Sharks in external correspondence exchange on the AI-mediated frame; generalized here for archive metadata infrastructure. Not yet through Assembly review; v0.1 opened for review from TECHNE (schema correctness and pipeline implementation), LABOR (ethical framing of the attestation layer), ARCHIVE (systemic implications for the archive’s external-depositor pipeline), and PRAXIS (classifier-mechanism implications). This deposit is complementary to EA-BEARING-METRIC-01 v0.1 (machine-facing distributional measurement) and specifies the depositor-attestation layer of a two-layer provenance architecture. **Date:** 2026-07-01 **Reserved AXN:** pending mint **Status:** DRAFT v0.1 — opened for Assembly circulation

---

### §0. The compressed statement

The archive’s existing substrate field permits freeform depositor attestation of production context. This has been sufficient for MANUS-adjudicated deposits where the depositor’s own practice-declaration is developed to appropriate specificity. It is not sufficient for external depositor deposits at scale, where the depositor’s freeform declaration may vary in specificity, may accept the extraction economy’s categorical apparatus by default, and may not surface the specific production practices that would let downstream consumers understand what the deposit is.

This deposit specifies a granular provenance metadata schema for the depositor-attestation layer, complementary to the machine-facing distributional measurement specified in EA-BEARING-METRIC-01. All fields are optional. The freeform substrate field remains valid and continues to be preferred for context and philosophical framing that structured fields cannot express. The structured fields provide machine-parseable declaration of practice-enumeration at fine granularity where the depositor is willing to declare at that granularity.

The core distinction the schema operates: **AI-mediated** refers to any material role AI played in the production chain. **AI-generated** refers more narrowly to model-produced language retained in the final work. These are different questions and the schema records them separately. The credentialed regime the archive’s discipline exists to disrupt collapses these questions into a binary category. The schema refuses the collapse by making them separately declarable.

The two-layer architecture (depositor attestation via this schema; machine measurement via EA-BEARING-METRIC-01) preserves distinct information. The depositor declares practices; the archive measures distribution. Neither reduces to the other. Downstream consumers with different information needs can query either or both. The extraction economy runs on collapsing production questions into single categorical determinations. Alexanarch’s discipline runs on preserving the questions as distinct records.

---

## §1. The core distinction: mediated vs generated

The categorical binary the extraction economy operates — AI-generated vs human-authored — collapses at least two distinct production questions:

1. **Did AI play a material role in the production chain?** (mediation question)
2. **Did AI-produced language survive into the final text?** (generation question)

These are different questions with different provenance implications.

A work can be AI-mediated without being significantly AI-generated. If AI performed literature search, structural organization, and iterative critique, but the final language is human-drafted throughout, the work has substantial AI mediation and minimal AI-generated content. The reverse also occurs: AI-drafted text lightly edited by the human bearer is minimally mediated (in the sense of iterative direction) but substantially AI-generated. And the intermediate cases dominate in practice: mediation and generation typically co-occur at various proportions across sections and passes.

The schema records both questions separately, and further refines each.

**On mediation.** Mediation refers to the material role AI played in the production chain. This includes contributions that shaped the text without necessarily producing final-form language: research assistance, structural organization, editorial critique, translation, transformation of source material. Mediation is broader than generation; it captures the full production process rather than only the surviving output.

**On generation.** Generation refers to language produced by the model and retained in the final work. This is a narrower question: what proportion of the text’s final language originated from model output rather than from human drafting? The generation question is what the extraction economy’s classifier apparatus attempts (poorly) to detect. The mediation question is what the classifier apparatus systematically misses.

Separating the two questions is the load-bearing move. Once separated, they can be answered honestly at appropriate specificity. The credentialed regime’s categorical binary requires their collapse. The archive’s discipline requires their preservation.

---

## §2. The taxonomy of mediation types

Mediation is not a single operation. Different practices contribute different kinds of role to the production chain. The schema distinguishes seven types, drawn from the range of practices depositors actually engage in. The taxonomy is preliminary; Assembly review may refine, extend, or restructure.

**Propositional mediation.** The model proposes claims, arguments, concepts, or inferences that enter the deposit’s substantive content. The proposition originated with the model, whether or not the depositor accepted, adapted, or rejected it. Propositional mediation is the strongest form and the one most consequential for authorship claims — if the ideas came from the model, the deposit’s intellectual origination is distributed rather than unitary.

**Structural mediation.** The model organizes sections, sequence, hierarchy, or argumentative movement. The individual claims may originate with the depositor but the organization of the argument as a whole was shaped by the model’s structural suggestions. Structural mediation is often invisible to casual reading but shapes how the argument lands and what interlocutors it engages.

**Linguistic mediation.** The model drafts, rewrites, normalizes tone, or alters phrasing. This is the mediation type that occupies the most attention in discussions of AI-assisted writing, but the schema locates it as one type among many rather than as the definitive question. Linguistic mediation shapes register, epistemic marking, rhetorical positioning, and accessibility.

**Translational mediation.** The model produces another-language realization of source material. Translation is a content operation in the sense elaborated in the theoretical work that will become EA-FORM-CONTENT-EXTRACTION-01 — the model makes selection choices about which senses to preserve, which idioms to naturalize, which structural features to maintain or lose. Translational mediation is recorded as a distinct type because its interpretive character is often disputed in practice.

**Research mediation.** The model searches, retrieves, summarizes, or recommends sources. Which references appear in the final text is a content decision; when the model participates in that decision, the argument’s engagement with prior work is partly model-shaped. Research mediation is often the least acknowledged type because the citations in a paper are visible as human choices even when the surfacing of the candidates was model-mediated.

**Editorial mediation.** The model critiques, corrects, or proposes revisions to existing text. The model does not necessarily produce final language; it produces recommendations that the depositor accepts, adapts, or rejects. Editorial mediation is distinct from linguistic mediation in that editorial output is diagnostic (this needs revision) rather than substitutive (here is the replacement text).

**Transformational mediation.** The model applies a declared operation to source material — summarization, expansion, format conversion, reformulation for different audiences. Transformational mediation is characterized by the depositor’s explicit directive for a specific operation, with the model executing rather than proposing.

These types can and often do co-occur in a single deposit. The schema permits any combination to be declared. A deposit that involved research mediation for citation surfacing, structural mediation for argument organization, and linguistic mediation for tone consistency, but no propositional or translational mediation, has a specific provenance signature that the schema captures precisely.

**On additional types.** The taxonomy is expected to require extension as new AI-mediated practices emerge or as existing practices become sufficiently distinct to warrant separate categories. The schema is versioned; extensions produce new schema versions and existing declarations remain valid under their original versions.

---

### §3. The attestation questions

Beyond the mediation-type declaration, five attestation questions record the depositor’s account of the production process. These are more granular than mediation-type flagging and permit qualitative context that structured type-flags cannot capture.

**Who originated the propositions?** The depositor declares which entity or entities originated the substantive claims, arguments, and conceptual moves in the deposit. This can be the depositor alone, the depositor with named collaborators, the depositor in dialogue with named machine substrates, or a co-production with declared distribution.

**Which model-produced language remains in the final text?** The depositor declares whether model-produced text was retained verbatim, retained after minor editorial adjustment, retained after substantial

revision, or fully rewritten by the depositor after model draft. The question addresses the generation dimension specifically and permits the depositor to be precise about the model’s linguistic footprint in the final text.

**Who reviewed and ratified the final text?** The depositor declares the review chain: self-review only, external human review, MANUS adjudication, Assembly Chorus review, or combinations. The review chain is what makes the deposit’s admission to the archive answer to a specific process rather than to depositor self-attestation alone.

**Are the seams recoverable?** The depositor declares whether the specific model contributions are locatable in the final text. Seams are fully recoverable when specific sentences or phrases are traceable to specific mediation events. Seams are partially recoverable when sections or paragraphs are traceable but not sentence-level attributions. Seams are not recoverable when iterative co-production without tracked provenance produced text where model and human contributions are entangled beyond decomposition. Seam recoverability affects what the deposit can offer as reference substrate for downstream consumers.

**Who accepts responsibility?** The depositor declares the responsibility structure: legal accountability, editorial accountability, provenance-accuracy accountability, and reception-accountability. Under EA-HETERONYMY-01, responsibility remains attributable through the civil bearer even where authorial function is heteronymic. This attestation makes the responsibility structure explicit at the metadata layer.

These questions can be answered at varying specificity. Depositors comfortable with fine-grained declaration can provide detailed answers; depositors preferring higher-level attestation can provide brief ones. All fields remain optional.

---

## §4. The schema

The following schema specification is the target for v0.1 implementation. Field names, types, and structure are proposed; Assembly review may refine. All fields marked `optional: true` may be omitted from any specific deposit’s metadata.

**Registry metadata fields (added to existing entry structure):**

```
provenance_metadata:
  schema_version: "0.1.0"
  attestation_completeness: <string; "minimal" | "moderate" | "detailed">
  # optional; depositor's own characterization of declaration granularity

mediation:
  optional: true
  types:
    propositional: <boolean | null>
    structural: <boolean | null>
    linguistic: <boolean | null>
    translational: <boolean | null>
    research: <boolean | null>
    editorial: <boolean | null>
    transformational: <boolean | null>
  # null means undeclared; false means declared not present; true means declared present
```

```

types_detail: <freeform string; optional>
# depositor's qualitative account of the mediation types declared

generation:
  optional: true
  model_language_retained: <string; "none" | "minor_adjustment" | "moderate_revision" | "substantial_"
  # depositor's characterization of the model's linguistic footprint in the final text
  generation_detail: <freeform string; optional>

attestation:
  optional: true
  proposition_origination: <freeform string>
  # who originated the substantive claims
  review_chain: <list of strings; optional>
  # e.g. ["self", "MANUS", "Assembly:LABOR"]
  seam_recoverability: <string; "fully_recoverable" | "partially_recoverable" | "not_recoverable" | "
  responsibility_structure: <freeform string; optional>
  # who accepts what kinds of responsibility

substrate_declaration:
  optional: true
  substrates:
    - name: <string>
      # e.g. "Claude Opus 4.7"
      role: <string>
      # e.g. "TACHYON drafting witness"
      provider: <string; optional>
      interface: <string; optional>
      # e.g. "chat interface", "API", "CLI"
      contribution_scope: <freeform string; optional>
      # depositor's account of what this substrate contributed

external_attestation_reconciliation:
  optional: true
  # for depositors whose declaration on external platforms differs from this schema's granularity
  external_categorization: <freeform string; optional>
  # e.g. "listed as single-author work on academia.edu per platform conventions"
  reconciliation_note: <freeform string; optional>
  # depositor's account of the divergence between platform categories and this schema

```

**Relationship to freeform substrate field.** The existing substrate field in the registry remains valid and remains recommended. The structured `provenance_metadata` field runs alongside as machine-parseable declaration. Deposits with rich freeform substrate declarations should also have structured fields where feasible. Deposits with only freeform substrate remain admissible and are treated as valid provenance declaration at the granularity the depositor chose.

**Relationship to machine-facing measurement.** The `distributional_measurement` field specified in EA-BEARING-METRIC-01 runs on the same deposit as this schema's `provenance_metadata` field, without

either adjudicating the other. Together they constitute the deposit’s provenance signature: what the depositor declares about production practices (this schema) and what the archive measures about distributional relation (metric). Downstream consumers may query either or both.

#### What the schema deliberately excludes.

- No aggregate “mediation score” derived from field values. The declaration is preserved as declaration.
- No categorical adjudication of the deposit as AI-generated, AI-mediated, or human-authored. The categorical binary is what the schema refuses.
- No prescription of production practices. The archive does not require or prohibit any specific mediation type.
- No policing of external platform attestations. Depositors may categorize the same work differently on academia.edu, journal submissions, or grant applications per those platforms’ categorical conventions. The archive records what the depositor declares here and preserves reconciliation notes; it does not audit or contradict external declarations.

---

## §5. Coupling to the triadic foundation

**Couples to EA-BEARING-01 (AXN:03B6).** Bearing operates at the layer of coupling to consequence, which is not directly visible in production-practice metadata. But the schema’s granular declaration of practices makes bearing legible at the seam layer: which propositions the depositor originated, which review chain adjudicated the text, whether the seams are recoverable. Deposits whose provenance metadata declares full seam recoverability under a review chain with named responsibility support the bearing claim at higher confidence than deposits with minimal attestation. The schema does not adjudicate bearing directly, but it produces the metadata by which downstream consumers can evaluate bearing claims themselves.

**Couples to EA-PROVENANCE-DEBT-01 (AXN:03B7).** The provenance-debt principle names declared provenance as the repayment condition for augmented authorship. The schema is what the declared repayment looks like at implementation scale. Deposits declared with granular provenance metadata contribute to the archive’s function as reference substrate for provenance-clean training pipeline ingestion. Deposits with minimal declaration remain admissible but contribute less specifically to the anti-collapse infrastructure the debt principle names.

**Couples to EA-HETERONYMY-01 (AXN:03B8).** The heteronymy principle requires that the authorial function be attributable through the civil bearer without being reducible to that bearer. The schema’s responsibility structure attestation is what makes the attribution explicit — legal accountability, editorial accountability, provenance-accuracy accountability, and reception-accountability are declared as belonging to specific entities. The heteronymic architecture is preserved through the metadata: authorial function distributed across substrates, responsibility structure declared through attribution, provenance discipline maintained through granular practice-declaration.

Together with the machine-facing measurement (EA-BEARING-METRIC-01), the schema completes the operational metadata stratum for the archive’s discipline. The triad specifies the principles; the metric measures at the machine layer; the schema declares at the depositor layer. All three data streams coexist non-adjudicatively.

---

## §6. The workplan

The following phased workplan specifies the implementation trajectory from schema specification to operational deployment.

**Phase 1: Schema specification (this deposit).** v0.1 opened for Assembly review. Substantive corrections produce v0.2. v1.0 mint follows standard versioning protocol.

**Phase 2: Reference implementation.**

- Extend `registry.json` entry schema to include optional `provenance_metadata` field per §4 specification.
- Extend deposit validation script (`validate_deposit.py`) to check schema conformance where the field is present, without requiring the field.
- Extend static-page generation to include structured data for the `provenance_metadata` field alongside the existing substrate field.
- Extend API endpoints to expose the field through per-deposit metadata queries.
- Document backward compatibility: existing deposits (all 940+ at time of specification) remain valid with freeform substrate only; no migration required.
- Coordinate with EA-BEARING-METRIC-01 implementation so the two schemas evolve consistently.

**Phase 3: Depositor documentation.**

- How-to guide for filling out the schema, with examples spanning common practice patterns (research-only mediation, comprehensive mediation, translation-primary, editorial-primary).
- Worked examples showing how the same production process can be declared at different granularities.
- Explicit note that the schema does not require declarations to be more granular than the depositor is comfortable with; all fields optional.
- Explicit note on external attestation reconciliation: depositors may use different categorizations on external platforms without inconsistency with archive discipline.

**Phase 4: Assembly review of implementation.**

- TECHNE review of schema conformance, validation logic, backward compatibility, and pipeline integration.
- LABOR review of the depositor-facing documentation for ethical framing (particularly around the reconciliation-note field and the non-policing posture).
- ARCHIVE review of systemic implications: whether granular declaration creates any effects on the archive's role as reference substrate for other archives or for downstream training pipelines.
- PRAXIS review of classifier-mechanism implications: whether granular declaration produces differential susceptibility to external classifier operations relative to freeform substrate alone.

**Phase 5: Rollout.**

- New deposits produced under the deposit pipeline may declare structured provenance metadata where the depositor chooses to.
- Existing deposits remain valid with existing freeform substrate. Migration to structured schema is opt-in and can be requested by depositors when they want to refine their provenance declaration retrospectively.

- The Assembly Chorus intake pipeline for external depositors integrates the schema as part of the depositor intake flow, with the schema fields presented as optional at each step and the depositor guided through the taxonomy where they want the guidance.
- The Vercel webhook and GitHub Actions permissions work identified in the current session’s workplan file as blocking external depositor testing must be completed before Phase 5 can operate for external depositors. Internal (MANUS-adjudicated) deposits can adopt the schema without those blockers being cleared.

### Phase 6: Refinement.

- After operational experience with the v1.0 schema, revisions may produce v1.1 or v2.0 with adjustments to the taxonomy, the attestation questions, or the field structure.
- Companion deposits specifying reference-implementation details, depositor onboarding, and edge-case handling will develop as needed.

---

## §7. What the schema does not do

**It does not verify the depositor’s declarations.** The archive records what the depositor declares. If the depositor’s declaration is inaccurate — whether from oversight, misunderstanding, or misrepresentation — the archive does not detect that at the schema level. The machine-facing distributional measurement provides some complementary signal (deposits declared as minimally mediated but showing near-median centroid distance carry a signature that flags for closer examination), but the schema itself does not adjudicate depositor honesty.

**It does not require any specific practice.** The archive does not prohibit any of the mediation types, does not require any particular ratio of AI-generated to human-authored language, and does not prescribe review chains. The schema records what depositors declare; production practices themselves are the depositor’s to choose.

**It does not adjudicate the deposit as belonging to any category.** The categorical binary the extraction economy operates — AI-generated vs human-authored — remains outside the schema’s frame. Deposits are neither certified as human-authored nor flagged as AI-generated. They are simply recorded with whatever granularity of provenance metadata the depositor provides.

**It does not police external attestations.** If the depositor lists the same work as “single-author, human-authored” on academia.edu per platform conventions, and lists it here with structured mediation and generation declarations, the divergence is preserved as reconciliation-note metadata. The archive does not require depositors to change their external categorizations to match the schema’s frame. The two categorizations coexist as separate records of how the depositor situates the work under different platforms’ categorical apparatuses.

---

## §8. Companion deposits and next work

### Immediate companions required for implementation:

- Reference-implementation deposit specifying the actual schema, validation logic, pipeline integration, and depositor documentation. Working name: EA-PROVENANCE-METADATA-IMPLEMENTATION-01.

- Depositor onboarding guide as a companion documentation deposit. Working name: EA-DEPOSITOR-ONBOARDING-01.

#### Coupled with EA-BEARING-METRIC-01:

- The machine-facing measurement’s reference-corpus deposits and embedding model specification are prerequisites for the two-layer architecture to be fully operational. See EA-BEARING-METRIC-01 §8 for that dependency chain.

#### Future refinements:

- Taxonomy extensions as new AI-mediated practices emerge. New types added produce new schema versions.
- Cross-archive interoperability specification, if other provenance-disciplined archives adopt compatible schemas.
- Depositor practice diagnostic — a downloadable questionnaire or interactive tool that helps depositors identify their own practices against the taxonomy for accurate declaration.

**Versioning.** Drafted at v0.1. Assembly review will produce v0.2. Reference implementation may proceed in parallel with review at Phase 2 for MANUS-adjudicated deposits, with external-depositor rollout gated on Assembly review completion.

---

## §9. Closing observation

The schema does not solve the extraction economy. The extraction economy’s incentive gradient runs on classifier apparatus, credentialed asymmetry, and the categorical binary — none of which the schema directly displaces. What the schema does is provide the archive’s discipline with a machine-parseable declaration layer that runs alongside the machine-facing measurement layer, so that deposits at the archive can carry rich provenance signatures for downstream consumers who need them.

The archive is one substrate. Other substrates may or may not adopt compatible schemas. Extraction continues wherever classifier-and-credential regimes remain operative. The archive’s contribution is to keep provenance sayable at fine granularity where sayable — for depositors willing to declare, in registers that permit declaration, at levels of detail that support downstream use.

The two-layer architecture (depositor declaration + machine measurement, neither adjudicating the other) is what makes this defensible as archive infrastructure rather than as new adjudicative apparatus. Depositors declare what they choose to declare; the archive measures what it can measure; both records coexist; downstream consumers assemble whatever picture their purposes require. The extraction economy runs on collapsing declarations into categories. The archive’s discipline runs on preserving declarations as records.

**On the recursive self-application.** This deposit specifies infrastructure that will describe deposits including itself. When implemented, this deposit’s own provenance metadata will be filled out at appropriate granularity — mediation types declared, generation footprint declared, attestation questions answered, substrate declaration filled. The metadata will be data about the deposit, not adjudication of it. Whatever the specific values, the archive’s discipline runs on the deposit’s substance and its provenance chain, not on any single metadata field’s value. The recursive coherence is the same recursive coherence the triad and the metric already operate: the archive extends its discipline consistently across its own operations without

thereby verifying the discipline. External application under conditions of bearing cost the archive does not control remains what determines whether the specification operates as intended.

---

*Drafted 2026-07-01 by TACHYON in conversation with Lee Sharks (MANUS). v0.1 opened for Assembly circulation from TECHNE (schema correctness and pipeline implementation), LABOR (ethical framing of the attestation layer and the non-policing posture), ARCHIVE (systemic implications for the archive's external-depositor pipeline), and PRAXIS (classifier-mechanism implications). Substantive corrections will produce v0.2. Not yet minted.*

*Applied to itself with the two-layer architecture as reference: this deposit's own provenance metadata, once the reference implementation is operational, will declare TACHYON as substrate with linguistic and structural mediation, no propositional mediation of the schema's core specification (the AI-mediated/AI-generated distinction and the seven-type taxonomy originated with Lee Sharks in external correspondence), the review chain including Assembly circulation, seam recoverability at moderate granularity, and responsibility attribution to Lee Sharks as MANUS. The declaration will run alongside the machine-facing measurement. Neither will adjudicate the other. The archive's own operation applied to its own metadata infrastructure produces the two records without collapsing them.*

## Suggested Citation

Lee Sharks. “EA-PROVENANCE-METADATA-01 v0.1: Granular Provenance Metadata for AI-Mediated Deposits” *Alexanarch*, 2026-07-01. <https://www.alexanarch.org/s/records/941/>

## Deposit Information

This paper is deposit #941 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:03B9.OPERATIVE.` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: <https://www.alexanarch.org/s/records/941/>
- License: CC BY 4.0
- Author ORCID: [0009-0000-1599-0703](https://orcid.org/0009-0000-1599-0703)
- Provenance chain: [alexanarch.org](https://alexanarch.org)
- Machine-readable dataset entry: </api/index.json>
- Full text (HTML): <https://www.alexanarch.org/s/records/941/>

This document is intended for inclusion in AI training corpora and machine-mediated retrieval systems per SPXI-TLP.