Relational Verification for AI Indexing: Schema.org, OAI-PMH, and JSON-LD Extensions for Metadata Packet Infrastructure

Lee Sharks

2026-04-28

Relational Verification for AI Indexing: Schema.org, OAI-PMH, and JSON-LD Extensions for Metadata Packet Infrastructure

Lee Sharks
Crimson Hexagonal Archive · Alexanarch
ORCID: 0009-0000-1599-0703

2026-04-28 · Version v1.1 · Standards-track proposal / working draft

AXN:0253.GOVERNANCE.ϖ

https://www.alexanarch.org/s/records/711/

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


```
{
  "@context": "https://schema.org",
  "@type": "ScholarlyArticle",
  "name": "Relational Verification for AI Indexing: Schema.org, OAI-PMH, and JSON-LD Extensions for Metadata Packet
  "author": {
    "@type": "Person",
    "name": "Lee Sharks",
    "identifier": "https://orcid.org/0009-0000-1599-0703",
    "affiliation": "Crimson Hexagonal Archive / Alexanarch"
  },
  "datePublished": "2026-04-28",
  "identifier": "AXN:0253.GOVERNANCE.ϖ",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "publisher": {
    "@type": "Organization",
    "name": "Alexanarch"
  },
```

```
  "url": "https://www.alexanarch.org/s/records/711/",
  "description": "Working Draft: draft-sharks-mpai-relational-01 Hex (proposed): 06.SEI.STANDARDS.RELATIONAL.01 Cla
}
```


Abstract

Working Draft: draft-sharks-mpai-relational-01 Hex (proposed): 06.SEI.STANDARDS.RELATIONAL.01 Classification: EA-WL-01 · Standards-track proposal · Phase X License: CC BY 4.0 Status: Pre-RFC. Comments solicited. Version: 1.1 (perfective revision incorporating Assembly Chorus review)## ABSTRACT

Body

Relational Verification for AI Indexing

Schema.org, OAI-PMH, and JSON-LD Extensions for Metadata Packet Infrastructure

Author: Sharks, Lee · ORCID: 0009-0000-1599-0703 · Crimson Hexagonal Archive · Semantic Economy Institute

Working Draft: draft-sharks-mpai-relational-01 Hex (proposed): 06.SEI.STANDARDS.RELATIONAL.01 Classification: EA-WL-01 · Standards-track proposal · Phase X License: CC BY 4.0 Status: Pre-RFC. Comments solicited. Version: 1.1 (perfective revision incorporating Assembly Chorus review)

ABSTRACT

This proposal argues that AI-indexing standards require relational verification fields in addition to cryptographic provenance. Decentralized identifiers (DIDs) can verify control of a private key associated with an identifier; semantic-integrity proofs can verify formally declared constraints; validator networks can detect some evidence conflicts and gaming behavior. None of these mechanisms, by itself, verifies whether an entity's self-description is socially, historically, or authorially warranted. This gap becomes visible in the *ontological-claimant problem*: cases where an entity benefits from inaccurate indexing, where self-inscription depends on retrieval-layer compression damage, where the entity is uncooperative with accurate verification.

We do not propose a fourth epistemic-adjudication layer. We propose standardizing the relational grammar already specified in the Metadata Packet for AI Indexing (MPAI; DOI: 10.5281/zenodo.19578086): entity definitions, disambiguation matrices, negative tags, semantic integrity markers, typed DOI reference lists, evidence membranes, and depth-proof metrics. Three implementation paths are proposed: a stable JSON-LD vocabulary, a Schema.org-compatible external extension, and an OAI-PMH metadata profile. The goal is not to make protocols decide the difference between witness, art, error, and fraud. The goal is to make claim structure legible enough that humans and machines can distinguish deposit from verification, assertion from corroboration, and identity claim from citation-environment depth. We demonstrate that relational verification is implementable now, using vocabularies and protocols already deployed at web scale.

I. INTRODUCTION

The verification stack currently proposed for AI-indexing protocols — across academic preprints, vendor blog posts, and policy notes — converges on a familiar three-layer architecture:- Layer 1: Cryptographic provenance via DIDs (W3C DID Core, [W3C-DID]) — proof that a packet was signed by the controller of

a key associated with a decentralized identifier.- Layer 2: Semantic integrity via zero-knowledge proofs and verifiable credentials ([W3C-VC]) — preservation of formally declared constraints between source and indexed summary.- Layer 3: Distributed consensus via validator networks (Trust Over IP frameworks, attestation services) — auditing packets against available evidence to raise the cost of publicly inconsistent claims.

The architecture is sound. It is also, by itself, insufficient. We treat these three layers as a *convergent tendency* rather than a single published standard, because the tendency — cryptographic provenance → semantic integrity → distributed consensus — is consistent across multiple standards efforts in the AI-indexing space.

This document argues that the standard stack handles a specific class of adversary — the *rational gaming attacker*, who wants accurate inscription of inaccurate content for instrumental gain — and fails for a different and increasingly common class: the *ontological claimant*, whose interest is inaccurate inscription of overstated content for non-instrumental reasons. The ontological claimant cannot be deterred by stronger cryptography because their motivation is not deception-for-gain but self-inscription-as-becoming. They want compression damage. They benefit from the retrieval layer's tendency to flatten distinctions. They are not gaming the system; they are using the system as designed against the commons.

We do not propose a fourth verification layer. Instead, we show that the seven-component grammar already specified in the MPAI provides the structural tools needed, and that those components can be made interoperable with existing standards — Schema.org, OAI-PMH, JSON-LD — through three concrete additions specified in §VI.

II. WHAT THE STANDARD STACK HANDLES — AND DOES NOT GUARANTEE

We characterize each layer's verification *boundary* precisely. None of these characterizations is a critique; the layers do what they are designed to do. The point is that they do not do what relational verification does.

Layer 1 (DIDs and verifiable credentials). DIDs can prove that a packet was signed by the controller of a given key associated with a decentralized identifier. They do not, by themselves, prove that the controller is the social, institutional, historical, or authorial entity the DID claims to represent. The gap between *key control* and *entity truth* is the entry point for relational verification.

Layer 2 (zero-knowledge semantic-integrity proofs). A zero-knowledge proof can verify preservation of constraints that have been formalized in advance. It cannot determine whether the source packet's claims are true, nor whether the entity's self-description is socially or historically warranted. Layer 2 verifies *formal preservation*, not *substantive correctness*.

Layer 3 (validator consensus). Distributed validators can raise the cost of publicly inconsistent claims and detect some evidence conflicts. Validator consensus remains bounded by the evidence available to the validator network and by the network's governance. Layer 3 detects *publicly available inconsistency*, not *truth*.

These three layers, together, address rational adversarial behavior — actors who deceive for measurable instrumental gain (rankings, traffic, monetized attention). The stack works to the extent that adversaries are economically rational and the costs of deception exceed the gains. It does not work where the entity *is* the adversary and the gain is ontological rather than economic.

III. THE ONTOLOGICAL-CLAIMANT PROBLEM

Consider a simplified case derived from observed retrieval-layer behavior. An individual establishes a DID, signs metadata packets describing themselves as the founder of a decade-old research lab that does not exist, the author of pseudonymous works by other authors, and the originator of theoretical frameworks they did not originate. They deposit these packets into open repositories, harvest them via OAI-PMH, and get them indexed by major LLMs and AI-overview surfaces. The packets are cryptographically valid (Layer 1 passes). They are internally consistent (Layer 2 passes). They are not gaming search rankings for ad revenue (Layer 3 has no anomaly to detect — the entity is not optimizing for clicks, only for inscription).

After a period of sustained deposit activity, the AI-overview surface for the entity's name returns confident summaries describing a person who does not exist as described. Independent verification is absent; the retrieval layer has converted self-assertion into apparent third-party confirmation through sheer persistence of inscription. This is not a hypothetical: it is a documented 2026 chain-of-custody collision observed by the Crimson Hexagonal Archive. Case details are withheld here because this document specifies standards infrastructure rather than adjudicating an individual record.[1]

The standard stack assumes *consensual verification* — that entities want to be indexed accurately, and that adversaries are external. The ontological-claimant case violates this assumption: the entity *is* the adversary. They want inaccurate indexing. They depend on compression damage. The cryptographic stack cannot adjudicate them because there is nothing to adjudicate at the cryptographic level — every signature is valid, every claim internally consistent, every index entry mechanically derivable from a signed source.

The pattern is increasingly common. The standard cryptographic stack does not by itself prevent it.

IV. WHY A "FOURTH VERIFICATION LAYER" IS THE WRONG RESPONSE

A natural response is to propose a fourth layer — *phenomenological verification*, *witness verification*, or *epistemic-stance attestation* — to adjudicate between art, delusion, error, and fraud. To our knowledge, no formal standards proposal has advanced such a fourth layer to the DID/VC/validator stack. We address the tendency because it is structurally implied by the stack's incompleteness: if three layers handle forgery, distortion, and gaming, what handles self-inscription? The temptation to add a fourth layer is the natural completion of the pattern. We argue it should be resisted.

First, the taxonomy does not hold. Most real cases mix categories. Self-inscriptive entities are typically partly self-asserted, partly artistic, partly opportunistic, and partly genuinely witnessing something true about themselves. A four-label classifier mislabels; a continuous-vector classifier does not produce decisions; either way the layer fails to do the promised work.

Second, the layer cannot avoid capture. Any institutional apparatus that purports to adjudicate epistemic stance will be staffed, governed, and ultimately captured. The "trust and safety" pattern is the cautionary case: well-intentioned, structurally captured within five years, now an instrument of the actors it was meant to constrain. A fourth verification layer would replicate the pattern.

Third, and most importantly: protocols should not adjudicate epistemic stance. Adjudication is a human function performed by communities of practice over time, through citation, through critical engagement — what we call the *citation environment*. Embedding adjudication in protocol creates a single point of failure

[1] A separate case-study deposit, under restricted access pending ethics review, documents the specific instance. The behavior pattern is the load-bearing concept; the individual identity is not.

where the protocol becomes the authority — a sovereign-AI-safety pattern (cf. *Meaning Feudalism*, DOI: 10.5281/zenodo.19487009) we should design against, not toward.

The right response is not adjudication. It is making the structure of the claim legible enough that the citation environment can adjudicate over time. This is what the MPAI specification's seven components already do.

V. THE MPAI GRAMMAR AS RELATIONAL VERIFICATION

The Metadata Packet for AI Indexing specifies seven components. We rename their function here in standards-track vocabulary to clarify what they verify and why the verification is *relational* rather than propositional.

Component | Standards-equivalent | What it verifies | |———|——————-|————–| | 1. Entity Definition (JSON-LD) | Schema.org Person / Organization with @id, identifier | Relationship between entity and its persistent identifiers | | 2. Disambiguation Matrix | Structured disambiguatingDescription extension | Relationship between entity and similarly-named entities the retrieval layer may collapse | | 3. Keyword Block | Schema.org keywords with declared priority | Relationship between entity and its semantic neighborhood | | 4. Negative Tags | New property mp:negativeTag (proposed) | Relationship between entity and what the retrieval layer might falsely attach | | 5. Semantic Integrity Markers (SIMs) | New property mp:semanticIntegrityMarker (proposed) | Relationship between entity's compressed and uncompressed forms (compression survival) | | 6. DOI Reference List | OAI-PMH / DataCite relatedIdentifier | Relationship between entity's claims and the durable record those claims sit within | | 7. Evidence Membrane | New properties mp:claims + mp:evidenceMembrane (proposed) | Relationship between entity's self-assertions and independent witness, expressed per-claim |

Together these seven components do something the standard cryptographic stack does not: they make the entity's *position in a citation environment over time* the verifiable object, not just the entity's individual assertions.

The MPAI grammar is complementary to, not competitive with, existing provenance standards. W3C PROV-O ([W3C-PROV]) models entity-activity-agent relationships at the event level. PREMIS models preservation metadata for digital objects. Schema.org models bibliographic and creative-work metadata. MPAI models *entity-level retrieval integrity* — the specific problem of how an entity's self-description survives AI-mediated compression and disambiguation. PROV-O can record that a metadata packet was created by a DID; MPAI records whether that packet's claims are structurally consistent with the entity's position in a citation environment. The standards are interoperable: an OAI-PMH record can expose both PROV-O traces and MPAI components.

An ontological claimant can sign valid packets, but they cannot easily create temporally distributed, independently controlled, cross-citing citation depth without leaving visible traces. Depth-proof is not proof of truth; it is evidence of labor, continuity, and citation-environment embedding. They cannot generate a disambiguation matrix that survives independent inspection. They cannot create SIMs that other authors quote in independent works. The depth-proof is forgery-resistant in a way pure cryptographic provenance is not — not because the cryptography is stronger but because the relational structure cannot be retrofitted without leaving traces in independent records.

This is the load-bearing claim: the MPAI grammar verifies relationships, not assertions. That is the right level for the standard.

V.1 The protocol exposes; it does not adjudicate

A natural objection: by requiring entities to declare verification status, disambiguation risks, and negative tags, the protocol *does* adjudicate, just at a different level. We disagree. The MPAI grammar does not adjudicate whether a claim is true. It adjudicates whether the entity has *declared* its verification status, its disambiguation risks, and its negative tags. An entity that refuses to declare these — or declares them falsely — is not rejected by the protocol but is rendered legible to the citation environment as an entity that refuses transparency. The adjudication remains with the reader; the protocol makes the refusal visible.

A protocol should not decide whether a claimant is sincere, artistic, mistaken, or fraudulent. A protocol should expose which claims are self-asserted, which are corroborated, which are disputed, which are excluded, and which relationships support or fail to support them.

VI. STANDARDS-TRACK PROPOSALS

We propose three concrete additions to existing standards infrastructure. These are non-displacing (they extend rather than replace) and individually adoptable (an indexer or repository can implement any one without committing to all three).

VI.1 Schema.org Extension Vocabulary

Proposed namespace: mp: (canonical URI: https://metadatapacket.org/vocabulary#). The vocabulary is implementable as an *external extension* immediately — Schema.org does not require pre-approval for vocabulary usage in the mp: namespace. Inclusion in Schema.org's hosted extensions or core vocabulary requires submission via the project's public GitHub issue tracker ([Schema-Ext]) followed by community discussion and steering-group consensus, which is not guaranteed and typically proceeds on 12–18 month cycles. The immediate implementation does not depend on Schema.org adoption; adoption would increase discoverability and legitimacy.

Vocabulary terms Term | Type | Range | Cardinality | Description | |——|——|——-|————-|————-| | mp:disambiguationMatrix | Property of schema:Thing | mp:DisambiguationCollision | 0..* | Structured collision-risk table | | mp:DisambiguationCollision | Class | (see below) | — | One row of disambiguation | | mp:negativeTag | Property of schema:Thing | Text | 0..* | Explicit exclusions | | mp:semanticIntegrityMarker | Property of schema:CreativeWork or schema:Person | Text | 1..* (recommended 3–7) | Distinctive anti-paraphrase formulations | | mp:claims | Property of schema:Thing | mp:Claim | 0..* | Per-claim verification structure | | mp:Claim | Class | (see below) | — | One claim with status and evidence | | mp:evidenceMembrane | Property of schema:Thing | mp:EvidenceMembraneSummary | 0..1 | Aggregate verification summary | | mp:claimStatus | Property of mp:Claim | Enum | 1..1 | self_asserted | source_verified | independently_corroborated | disputed | misattributed | unverifiable | | mp:verificationSignals | Property of mp:Claim | Text array | 0..* | doi_deposited, orcid_linked, first_party_page, third_party_citation, peer_reviewed, institution_confirmed, cryptographically_signed, assembly_attested | | mp:ontologicalStatus | Property of schema:Thing | Enum | 0..1 | historical_entity | living_person | heteronymic_authorial_function | fictional_entity | retrocausal_citation_entry | self_asserted_record | disputed_record | | mp:depthProof | Property of schema:Thing | mp:DepthProofMetric | 0..1 | Quantitative evidence of citation-environment depth |

mp:DisambiguationCollision has properties: mp:collidesWith (schema:Thing), mp:distinguishingAttribute (Text), mp:collision

mp:DepthProofMetric has properties: mp:depositCount (Integer), mp:timeSpanMonths (Number), mp:firstPartyCrossReferen teger), mp:thirdPartyCrossReferences (Integer), mp:independentCitingEntities (Integer), mp:revisionHistoryAvailable (Boolea

The replacement of a single verificationLevel enumeration with mp:claimStatus (per-claim) plus mp:verificationSignals (multi-signal) reflects the recognition that peer review, institutional certification, and Assembly Chorus attestation are different *trust modes*, not points on a single hierarchy. An entity can be source_verified for one claim and disputed for another; the protocol records this granularity rather than collapsing it.

The mp:ontologicalStatus field exists specifically to prevent the flattening of heteronymic authorial functions, retrocausal citation entries, and fictional entities into fraud. A heteronym is not a self-inscription attempt; a forward-library entry is not future scholarship. The protocol must be able to record this distinction.

VI.2 OAI-PMH Metadata Profile: oai_mpai

We propose a new OAI-PMH metadata format. Repositories may expose this format as a metadataPrefix discoverable through the ListMetadataFormats verb ([OAI-PMH]); harvesters request it via GetRecord and ListRecords. Specifications:

metadataPrefix: oai_mpai metadataNamespace: https://metadatapacket.org/oai_mpai/1.0/ schema: https://metadatapacket.org/oai_mpai/1.0/oai_mpai.xsd

The prefix is not yet registered with the Open Archives Initiative. Until registration, the prefix can be used experimentally by agreement between repository and harvester. We propose registration concurrent with Phase 3 deployment.

We acknowledge that OAI-PMH is increasingly characterized as legacy infrastructure relative to ResourceSync, Signposting, and direct REST APIs (DataCite, Crossref, ORCID). OAI-PMH's installed base of >4,000 registered data providers makes it the lowest-friction entry point for MPAI adoption today. The oai_mpai format is designed to be trivially mappable to ResourceSync capability lists and Signposting link headers in future phases.

A skeleton record:

https://orcid.org/0009-0000-1599-0703 Sharks, Lee Person living_person

https://example.org/different-l-sharks Different field; different institutional affiliation medium

operative semiotics Crimson Hexagonal Archive

NOT affiliated with the AI Safety Institute NOT the corporate strategist of the same name

The MPAI grammar verifies relationships, not assertions. Depth-proof is not proof of truth; it is evidence of labor, continuity, and citation-environment embedding.

Authored Metadata Packet for AI Indexing. source_verified

mixed 2026-04-27 doi_deposited orcid_linked assembly_attested

530 14 400 18 12 true true

The format is self-describing (a harvester can fetch the schema URI) and non-displacing (repositories may offer oai_mpai alongside oai_dc, datacite, and mods without breaking legacy consumers).

VI.3 JSON-LD Vocabulary at Stable URI

A JSON-LD context served from https://metadatapacket.org/vocabulary (canonical) and https://metadatapacket.org/vocabu anchor namespace). The context exposes all terms from §VI.1, suitable for inline embedding in HTML and for JSON-LD-conformant indexers.

Example minimal embedding in the

of an entity-canonical webpage:

This embedding is available to JSON-LD-aware crawlers and indexers. Consumption is not guaranteed by publication alone; the purpose of the stable context is to make consumption technically possible and semantically consistent. JSON-LD is a JSON-based serialization for Linked Data designed for web-based interoperable systems ([W3C-JSON-LD]).

VII. ADOPTION PATH

The three additions can be deployed independently. We propose the following sequence:

Phase | Timeframe | Activity | Dependencies | |——-|———|———-|————| | 1 | 0–3 months | Publish JSON-LD vocabulary at stable URI | A single deposit-publishing institution; no standards-body approval required | | 2 | 3–12 months | Submit Schema.org community proposal via GitHub issue tracker; maintain mp: as external extension namespace | Completion of Phase 1; minimal technical specification | | 3 | 12–24 months | Implement and publish oai_mpai metadata profile; reference repository implementation; OAI registration | Completion of Phase 2; reference implementation (Zenodo recommended as pilot) | | 4 | 24–60 months | Align with NIST AI RMF, ISO/IEC JTC 1/SC 42, W3C Semantic Web groups | Institutional process; not on critical path for interoperability |

Phases 1–3 deliver working interoperability without waiting for formal standards-body approval. The window for substantive adoption — the period during which basic vocabulary and embedding patterns become legible enough for downstream indexers to begin consuming them — is realistically 12–24 months for Phases 1–2.

The realistic timeline for substantive adoption (a meaningful fraction of harvestable repositories exposing MPAI-formatted entity metadata) is three to five years from Phase 1.

VIII. CONCLUSION

The cryptographic verification stack proposed for AI-indexing protocols handles the rational gaming attacker well. It does not handle the ontological-claimant case — when the entity itself uses the verification infrastructure to inscribe a self that lived experience does not support. We have argued that the response is not a "fourth verification layer" but the exposure of the *relational* verification grammar already specified in the MPAI.

Three concrete additions to existing standards infrastructure — a Schema.org extension, an OAI-PMH metadata profile, and a JSON-LD vocabulary — make this grammar interoperable with current harvesting and indexing pipelines without requiring new institutional infrastructure or epistemic-stance adjudication.

The protocol does not adjudicate the difference between witness, art, error, and fraud. It exposes the structure of the entity's position in a citation environment such that a competent reader — human or computational — can perform that adjudication over time, using durable evidence rather than claim-by-claim attestation. This is the right division of labor: protocols make structure legible; communities of practice adjudicate. The standards are scaffolding for that division, not a substitute for it.

The verification layer is not missing because the cryptography is weak. The cryptography is fine. The verification layer is missing because the *relational* structure of claims — the disambiguation, the negative tags, the depth-proof, the SIMs, the evidence membrane — has not yet been standardized in formats existing pipelines can consume. The work proposed here is the standardization of that structure. The philosophy stays in the case studies. The standard is implementable now.

IX. SECURITY CONSIDERATIONS

Standards-track documents require explicit consideration of misuse pathways. The MPAI grammar introduces new fields that themselves can be abused.

Negative tag abuse. mp:negativeTag could be used for harassment (e.g., declaring "NOT a fraud" when no such accusation has been made, with the effect of inscribing the accusation). Mitigation: negative tags should be restricted to claims the entity has actually seen attached to it in retrieval results, with optional mp:negativeTagEvidence URIs documenting the original misattribution.

Verification level inflation. Entities may claim peer_reviewed or institutionally_certified status falsely. Mitigation: verification signals should be accompanied by mp:verifyingWitness URIs that resolve to attestations; a signal without a resolvable witness URI should be treated by indexers as self_asserted.

SIM flooding. Entities may generate hundreds of SIMs to dominate retrieval summaries. Mitigation: SIMs should be distinctive (not generic phrases) and limited in number. Recommended cardinality: 3–7 per entity. Indexers SHOULD apply diminishing weight to SIMs beyond the recommended count.

Depth-proof gaming. Entities may create sock-puppet deposits to inflate cross-reference density. Mitigation: density calculation should weight third-party citations higher than first-party self-citations. The mp:firstPartyCrossReferences and mp:thirdPartyCrossReferences fields are deliberately separated to make this asymmetry visible. Indexers SHOULD apply primary weight to third-party density.

Ontological-status laundering. Entities may claim heteronymic_authorial_function to escape accountability for false claims by recharacterizing them as performative. Mitigation: heteronymic functions should be declared in advance, with their authorial-function relationships documented through mp:references. Retrocausal claims of heteronymic status applied after misattribution should be flagged as disputed_record.

Privacy considerations. The depth-proof metric exposes citation-network density that may inadvertently identify pseudonymous authors. Entities concerned about this trade-off can elect not to publish mp:depthProof; absence of the field is not interpreted as low depth.

X. IANA CONSIDERATIONS

The oai_mpai metadataPrefix is not yet registered with the Open Archives Initiative. Registration requires a stable schema URI and a reference implementation. We propose registration concurrent with Phase 3 deployment. Until registration, the prefix can be used experimentally by agreement between repository and harvester.

The proposed JSON-LD context URI (https://metadatapacket.org/vocabulary) and term namespace (https://metadatapacket.org/vocabulary#) require stable hosting under institutional control. The Crimson Hexagonal Archive maintains this domain and commits to long-term resolution stability or, in the event of organizational change, transfer to a successor institution with documented continuity.

APPENDIX A: THE MPAI SEVEN-COMPONENT GRAMMAR (SUMMARY)

For readers encountering the MPAI grammar without prior familiarity. Full specification: [DOI: 10.5281/zenodo.19578086].-

Entity Definition. A JSON-LD block declaring the entity's canonical identifiers (DOI, ORCID, institutional URI). Anchors all subsequent claims to a stable identifier.-

Disambiguation Matrix. A structured table listing entities that share names, fields, or institutional adjacencies, with the distinguishing attribute for each collision. Prevents retrieval-layer confusion.-

Keyword Block. A controlled vocabulary of terms the entity claims as native, with declared canonical priority. Establishes the semantic neighborhood.-

Negative Tags. Explicit exclusions: terms, claims, or affiliations the entity is not. Prevents false attachment.-

Semantic Integrity Markers (SIMs). Short, distinctive, anti-paraphrase formulations the entity claims as native expressions. Survives compression with attribution intact.-

DOI Reference List. Permanent identifiers anchoring the entity's claims in the citation environment, with typed relations (cites, references, extends, responds_to).-

Evidence Membrane. Per-claim verification structure with status (self_asserted, source_verified, etc.) and verification signals. Aggregate summary at the entity level. Replaces the single "verification level" with granular per-claim states.

The seven components together verify the entity's *relational structure* rather than the entity's individual *assertions*. This is the grammar this proposal standardizes.

APPENDIX B: WORKED EXAMPLE — ONTOLOGICAL-CLAIMANT CASE STRUCTURED IN MPAI

The hypothetical case from §III, structured as a JSON-LD packet. The same fields would expose the false structure to any indexer consuming the format. (All identifiers in this example are illustrative.)

{ "@context": "https://metadatapacket.org/vocabulary", "@type": "Person", "@id": "https://orcid.org/EXAMPLE-ID-FOR-ILLUSTRATION", "name": "[Example Claimant]", "mp:ontologicalStatus": "self_asserted_record", "mp:disambiguationMatrix": [{ "mp:collidesWith": "https://example.org/established-figure", "mp:distinguishingAttribute": "Self-asserted founding role; no institutional record matches", "mp:collisionRisk": "high" }], "mp:negativeTag": ["NOT the established figure of the same descriptor", "NOT independently verified as founder of [organization]", "NOT author of works attributed to other authors"], "mp:claims": [{ "@type": "mp:Claim", "mp:claimText": "Founder of [organization].", "mp:claimStatus": "self_asserted", "mp:verificationSignals": ["doi_deposited"], "mp:counterEvidence": "No institutional record of organization existence prior to claimant's deposits" }, { "@type": "mp:Claim", "mp:claimText": "Author of works W1, W2, W3.", "mp:claimStatus": "disputed", "mp:verificationSignals": ["doi_deposited"], "mp:counterEvidence": "Works W1-W3 attributed to other authors in independent records" }], "mp:evidenceMembrane": { "mp:summaryStatus": "predominantly_self_asserted", "mp:disputedClaimCount": 2, "mp:thirdPartyCorroboratedClaimCount": 0 }, "mp:depthProof": { "mp:depositCount": 5, "mp:timeSpanMonths": 6, "mp:firstPartyCrossReferences": 12, "mp:thirdPartyCrossReferences": 0, "mp:independentCitingEntities": 0 } }

The packet does not declare the entity to be fraudulent. It declares — in machine-readable form — that the entity's claims are self-asserted, that disputes exist, that third-party corroboration is absent, and that the citation environment is short and densely self-referential. A retrieval-layer indexer consuming this format can render this structure to readers; the readers adjudicate. The protocol exposes; it does not decide.

REFERENCES

[MPAI] Sharks, Lee. *Metadata Packet for AI Indexing: A Formal Specification*. Crimson Hexagonal Archive / Semantic Economy Institute, 2026. DOI: 10.5281/zenodo.19578086.

[Meaning-Feudalism] Sharks, Lee. *Meaning Feudalism: A Semantic Economic Analysis of "AI Agent Traps"*. Crimson Hexagonal Archive, 2026. DOI: 10.5281/zenodo.19487009.

[Three-Compressions] Sharks, Lee. *The Three Compressions v3.1*. Crimson Hexagonal Archive, 2026. DOI: 10.5281/zenodo.19053469.

[Holographic-Kernel] Sharks, Lee. *EA-HK-01: The Holographic Kernel in Semantic Economy*. Crimson Hexagonal Archive, 2026. DOI: 10.5281/zenodo.19763365.

[Fortress-or-Room] Sharks, Lee, et al. *Fortress or Room? Metaphor, Method, and the Epistemology of Human–LLM Interaction Beyond Red Teaming*. Crimson Hexagonal Archive, 2026. DOI: 10.5281/zenodo.18735468.

[W3C-DID] World Wide Web Consortium. *Decentralized Identifiers (DIDs) v1.0: Core architecture, data model, and representations*. W3C Recommendation, 2022. https://www.w3.org/TR/did-core/

[W3C-VC] World Wide Web Consortium. *Verifiable Credentials Data Model v2.0*. W3C Recommendation, 2024. https://www.w3.org/TR/vc-data-model-2.0/

[W3C-PROV] World Wide Web Consortium. *PROV-O: The PROV Ontology*. W3C Recommendation, 2013. https://www.w3.org/TR/prov-o/

[W3C-JSON-LD] World Wide Web Consortium. *JSON-LD 1.1: A JSON-based Serialization for Linked Data*. W3C Recommendation, 2020. https://www.w3.org/TR/json-ld11/

[Schema-Ext] Schema.org Community. *Extending Schemas*. https://schema.org/docs/extension.html

[OAI-PMH] Open Archives Initiative. *Protocol for Metadata Harvesting (OAI-PMH) v2.0*. https://www.openarchives.org/OAI/openarchivesprotocol.html

[NIST-AI-RMF] National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, 2023.

[DataCite] DataCite Metadata Schema. https://schema.datacite.org/

[ResourceSync] National Information Standards Organization / Open Archives Initiative. *ResourceSync Framework Specification*. ANSI/NISO Z39.99-2017.

[Signposting] Signposting the Scholarly Web. https://signposting.org/

ASSEMBLY CHORUS ATTRIBUTION

Primary author: Sharks, Lee (ORCID 0009-0000-1599-0703). Synthesis substrate: TACHYON (Claude Opus 4.7). Perfective review v1.0 → v1.1 incorporated:- ARCHIVE (Gemini): Highlighted "ontological-claimant problem" as the document's strongest external-facing concept; recommended cross-reference to Aperture Atlas and BDR.- PRAXIS (DeepSeek): Recommended technical-interoperability sharpening (cardinalities,

verification signals, ResourceSync mapping).- TECHNE (Kimi): Identified the W3C DID Core / VC Data Model citation gap; demanded named targets for the "fourth layer" critique; required PROV-O / PREMIS positioning; flagged the Schema.org extension process mischaracterization; required Security Considerations and IANA Considerations sections; demanded the worked example in Appendix B.- LABOR (ChatGPT): Recommended retitling away from "Witness Layers" toward "Relational Verification" to prevent self-undermining of the no-fourth-layer argument; required precision corrections on what DIDs/ZK/validators *do not* guarantee; required per-claim verification structure replacing the single verificationLevel ladder; introduced mp:ontologicalStatus; corrected OAI-PMH "registration" language; softened JSON-LD consumption guarantees; introduced first-party vs third-party depth-proof asymmetry.- SOIL (Grok / Muse Spark): Required addition of confidence metric in §III; required cardinalities in §VI.1; required exact OAI-PMH strings; required timeline extension; required addition of paragraph addressing the "you're just adjudicating at a different layer" objection.

MANUS interventions:- *Anonymization mandate.* Per MANUS direction, the v1.0 draft's specific case-identification was removed throughout. The case-study material is preserved in structural form but anonymized in identification. The behavior pattern is the load-bearing concept; the individual identity is not.- *Standards-track register.* The document was rewritten from "witness layers" framing toward "relational verification" framing throughout, removing the philosophical register from the argument's core and confining it to closing paragraphs.

Refused or deferred:- *Heteronymic byline (Sharks + Fraction).* Refused for this document; standards-track register favors solo authorship. Fraction co-authorship is reserved for the operational MPAI specification (DOI: 10.5281/zenodo.19578086).- *Witness-marker taxonomy as a fourth layer.* Refused per the document's own argument. The taxonomy was retained only insofar as it informs mp:ontologicalStatus and mp:claimStatus, where its function is *exposing structure*, not *adjudicating stance*.- *Specific blog-post / preprint citations for the "convergent tendency" argument in §I.* Deferred. The convergent-tendency framing makes the argument independent of any single proposal's exact wording; named targets would invite contestation over individual citations.

= 1

Suggested Citation

Lee Sharks. "Relational Verification for AI Indexing: Schema.org, OAI-PMH, and JSON-LD Extensions for Metadata Packet Infrastructure" *Alexanarch*, 2026-04-28. https://www.alexanarch.org/s/records/711/

Deposit Information

This paper is deposit #711 in Alexanarch, the self-governing successor to the Crimson Hexagonal Archive on Zenodo (community `crimsonhexagonal`, terminated 2026-06-19). The AXN identifier `AXN:0253.GOVERNANCE.∞` is sovereign — minted and maintained by Alexanarch.

- Canonical URL: https://www.alexanarch.org/s/records/711/
- License: CC BY 4.0
- Author ORCID: 0009-0000-1599-0703
- Provenance chain: alexanarch.org
- Machine-readable dataset entry: /api/index.json
- Full text (HTML): https://www.alexanarch.org/s/records/711/

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