The Interaction Claim Standard

Abstract

Independently implemented systems increasingly produce, exchange, and interpret claims about AI interactions: what participated, which interaction the supporting evidence concerns, and how far an assertion reaches. Where no shared requirements govern these claims, agreement on their meaning across implementations remains incidental rather than normative—each implementation may interpret the same assertion differently.

This document defines the Interaction Claim Standard, a standard for expressing and interpreting claims about AI interactions across independently implemented systems. It standardizes four normative objects: claim semantics, interaction-specific evidentiary binding, claim boundaries, and normative interpretation requirements. Together, these establish cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries. The standard builds on the Interaction Boundary Constraint and the Identity paper, and inherits the Identity paper's claim taxonomy. It is additive: where no standardized claim is asserted, the conditions established by the Interaction Boundary Constraint and the Identity paper remain unchanged.

The Interaction Claim Standard specifies claims, not systems. It prescribes no implementation architecture, evidence-production mechanism, or communication protocol, and standardized-claim status does not establish the truth of the assertion. The standard defines the structure, meaning, and boundaries of the claim; conformance determines whether an implementation satisfies those requirements; reliance remains external to both.

Thesis

Interoperable claims about AI interactions require a shared standard that defines claim semantics, binds evidence to a specific interaction, establishes validity and scope boundaries, and specifies normative interpretation requirements. These requirements establish cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries.

The standard does not determine whether the evidence supporting a standardized claim is sufficient for a particular reliance decision. It defines the structure, meaning, and boundaries of the claim; conformance determines whether an implementation satisfies those requirements; reliance remains external to both.

Scope

This document defines the **Interaction Claim Standard**, a standard for expressing and interpreting claims about AI interactions across independently implemented systems. It standardizes four normative objects: claim semantics, interaction-specific evidentiary binding, claim boundaries, and normative interpretation requirements. Together, these

establish cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries.

The Interaction Claim Standard specifies claims, not systems. It does not prescribe an implementation architecture, evidence-production mechanism, or communication protocol, and it defines no claim types beyond those inherited from the Identity paper's taxonomy (Harber, 2026b). Standardized-claim status does not establish the truth of the assertion.

The standard applies when standardized claims are asserted or interpreted. It does not require an implementation to assert a standardized claim as a condition of interaction, nor does it determine whether an implementation may interact with a counterparty that asserts none. Information that does not satisfy the standard's claim requirements is not a standardized claim and must not be interpreted as one. Refusing standardized-claim status is a standards decision; refusing communication is a policy decision.

The standard does not determine whether the evidence supporting a standardized claim is sufficient for a particular reliance decision. That determination depends on decision context supplied externally at evaluation time and remains external to the standardized claim and its conformance assessment.

The Interaction Claim Standard builds on the Interaction Boundary Constraint (Harber, 2026a) and the Identity paper (Harber, 2026b) without replacing the conditions they establish. Where no standardized claim is asserted, those conditions remain unchanged: interpretation relies on inference, and residual uncertainty remains regarding what is participating. The standard is additive; it defines what can be agreed upon when standardized claims are present without changing the condition of interactions in which they are absent.

Conformance determines whether an implementation satisfies these requirements. Reliance remains external to both.

1. Problem Statement — Why Interoperable Interaction Claims Fail Without Shared Semantics, Evidence Binding, Claim Boundaries, and Interpretation Rules

This section is a failure decomposition: four distinct ways cross-implementation agreement can break, each closed by one element the standard supplies. Structuring it this way keeps the four elements non-arbitrary—each exists to address a specific agreement failure, not because four felt complete. Every failure below is a failure of cross-implementation agreement, which is what the standard exists to establish.

Take the Interaction Boundary Constraint as given (Harber, 2026a): because producing composition cannot be **established** from the interaction surface alone, claims about that composition must be explicitly expressed and bound rather than inferred. That precondition makes the interoperability problem necessary—inference permits implementations to reach different interpretations of the same interaction.

- **Semantic-agreement failure:** Two independently implemented systems exchange an assertion but do not agree on what it means. Without shared claim semantics, there is no assertion in common—only a representation each implementation may interpret differently.
- **Binding-agreement failure:** The meaning is agreed, but the systems do not agree on which interaction the claim refers to, or one admits generic provenance, configuration, or artifact-level evidence as though it were evidence about the interaction under evaluation. Without agreed binding, evidence floats free of the interaction it is supposed to support.
- **Boundary-agreement failure:** Meaning and referent are agreed, but the systems do not agree on when the claim applies or what it covers. A partial claim can then be interpreted as reaching beyond what was actually asserted. Without declared validity and scope boundaries, claims can expand across implementations beyond their stated reach.
- **Interpretation-agreement failure:** Even with the first three failures closed, agreement remains incidental rather than normative unless implementations are required to interpret the assertion, its evidentiary binding, and its stated boundaries according to shared rules. Normative interpretation requirements make that agreement obligatory rather than incidental.

These failures are distinct: resolving one does not resolve the others. Shared claim semantics do not establish which interaction a claim concerns; an agreed evidentiary binding does not establish the claim's stated boundaries; and defining semantics, binding, and boundaries does not by itself require implementations to interpret them according to shared rules. Each element of the standard therefore addresses a separate requirement for cross-implementation agreement.

Exclusion: disagreement over whether the evidence is sufficient for a particular reliance decision belongs outside the interoperability problem defined here. Every failure in this section concerns cross-implementation agreement on the meaning of an assertion, its binding to a specific interaction, or its stated boundaries—not the downstream judgment made from that evidence.

2. Normative Objects — What the Standard Actually Standardizes

This section states precisely what is in the standard and what is merely referenced by it. It defines four normative objects, each corresponding to one element of the thesis and closing one failure identified in §1. Together, these objects establish cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries.

- *Claim semantics*: the expressible claim types and what each asserts. The standard inherits the Identity paper's claim taxonomy (Harber, 2026b) and defines the propositional semantics of those inherited types without adding new claim types. It fixes what each claim type means and what an assertion of that type does and does not assert; it does not determine whether a particular claim is true. This closes semantic-agreement failure.

- *Interaction-specific evidentiary binding*: the requirements by which a claim and its supporting evidence are associated with a specific interaction, including what must be true for that association to count as a binding rather than merely related information. The binding declares a specific interaction referent and must distinguish that referent from other interaction referents. Component referents remain governed by the Identity paper's identity-basis machinery (Harber, 2026b); interaction referents are distinct and do not resolve to component identity bases. For purposes of the standard, standardized claims whose resolved interaction-referent values are equal concern the same declared interaction. This closes binding-agreement failure.
- *Claim boundaries*: the explicit limits of what an assertion reaches and the conditions under which it applies. A claim must state enough about those limits and conditions that independently implemented systems can distinguish what is asserted from what lies outside the assertion, including where a partial claim does not extend to completeness or where a particular assertion concerning one component, interaction, or claim type does not extend beyond its stated reach. This instance-level non-extension operates within the type-level semantics defined by the standardized claim semantics. The standard fixes the meaning of those declared boundaries; it does not extend a claim beyond them. This closes boundary-agreement failure.
- *Normative interpretation requirements*: the rules independently implemented systems are required to apply when interpreting the assertion, its evidentiary binding, and its stated boundaries. A conformant implementation must interpret each claim type according to its standardized proposition; interpret the claim at the declared interaction referent rather than splitting, merging, or substituting another; and interpret the assertion only within its stated claim boundaries. These requirements govern interpretation of the standardized claim and do not constrain how an implementation internally models the underlying activity. This closes interpretation-agreement failure.

Standardized Claim-Type Semantics

All standardized claim propositions are interpreted within their stated claim boundaries. The standardized claim-type value space is closed and consists of **identity, invocation, observable participation, completeness, and material contribution**.

- **Identity.** An identity claim asserts that an entity or component is distinguishable from others and that references to it, within the claim's stated boundaries, resolve to the same stable identity basis. It asserts continuity of reference; it does not, by itself, assert invocation, observable participation, completeness, material contribution, authenticity, or any property of behavior.
- **Invocation.** An invocation claim asserts that an initiating request, call, or signal was directed to an identified component as part of the declared interaction, independent of the delivery mechanism. It does not, by itself, assert receipt, acceptance, successful execution, return of information, observable participation, or material contribution.

- **Observable participation.** An observable-participation claim asserts that an output, result, or action attributable to an identified component entered the production of the declared interaction. It does not, by itself, assert that the component's output, result, or action affected the interaction that was produced. The type name is inherited from the Identity paper; observability pertains to the evidence supporting the assertion rather than adding a separate observer-relative condition to the proposition defined here.
- **Completeness.** A completeness claim asserts that, with respect to another standardized claim type whose proposition is asserted relative to the declared interaction, the identified set is exhaustive: no entity or component outside that set satisfies the specified proposition for the declared interaction. Completeness does not, by itself, assert that any particular member of the identified set satisfies that proposition. Its admissible predicates are invocation, observable participation, and material contribution; identity is not interaction-indexed, and completeness cannot serve as its own predicate.
- **Material contribution.** A material-contribution claim asserts that an identified component affected the interaction that was produced. It asserts effect on the produced interaction, not merely invocation or entry into its production. It does not, by itself, establish completeness, correctness, quality, causal responsibility, or any other property beyond the stated contribution assertion.

The claim types are distinct assertions. A standardized claim of one type must not be interpreted as asserting another claim type merely because the underlying circumstances could support both. Where more than one standardized claim type is asserted, each assertion retains its own semantics, evidentiary binding, and stated claim boundaries.

The standard therefore defines the structure and meaning of claims, their interaction-specific evidentiary bindings, their stated boundaries, and the requirements governing their interpretation. It does not define how evidence is produced, prescribe which implementation architecture generates a claim, or determine whether a particular claim is true. It standardizes expression and interpretation, not production, veracity, or architecture.

Conformance and reliance are referenced by these normative objects but remain distinct from them. Conformance determines whether an implementation satisfies the standard's requirements (§3). Reliance concerns whether the evidence supporting a claim is sufficient for a particular decision context and remains external to both the standardized claim and its conformance assessment (§4).

Canonical Example

Identified component(s) → specific interaction → evidence → standardized claim → stated claim boundaries

An orchestration system identifies a set of components {Model B, Model C, Tool 1}, each of which presents an identity signal that resolves to a stable identity basis established under the Identity paper (Harber, 2026b). Evidence concerning a specific

interaction, *J*, is bound both to the identified set and to *J*. A standardized claim then asserts that the bound set is complete with respect to observable participation in *J*, where **completeness** carries the claim semantics defined by the standard and inherits the non-implication relations of the Identity paper's taxonomy.

The claim is explicitly bounded: it asserts only that no component outside the identified set observably participated in *J*. It does not establish the correctness or quality of any component's outputs, material contribution by any member of the set to the result of *J*, causal responsibility, or any other property beyond the completeness assertion. A partial identification of participants cannot be interpreted as a completeness claim; completeness, if asserted, must be asserted as such.

Layering Made Visible

- **Identity — Who are the components?** Each member of the set resolves to a stable identity basis. The mechanism that establishes that basis is not prescribed.
- **Binding — Which interaction does the evidence concern?** The binding declares interaction referent *J* and associates the evidence with *J* and the identified set, rather than describing the components in general.
- **Completeness — What is being asserted?** The standardized claim asserts that the bound set is complete with respect to observable participation in *J*: no component outside the set observably participated. The completeness assertion does not imply material contribution or any other unstated claim type or property.
- **Claim boundaries — What is not asserted?** Correctness, material contribution, causal responsibility, and properties outside the standard's claim taxonomy remain outside the assertion. The claim also does not establish that any particular member observably participated; it asserts only that no observable participant lies outside the identified set. Completeness is not implied merely because one or more components have been identified or shown to have observably participated.

Interpretation Constraints

When a conformant implementation asserts this standardized claim, it must express the claim, its interaction-specific evidentiary binding, and its stated claim boundaries according to the standard. When the claim is presented to a receiving conformant implementation, that implementation must interpret the claim at declared interaction referent *J* and must not extend the completeness assertion to material contribution, correctness, or any other unstated property.

Whether the evidence is sufficient for a particular downstream decision remains external to the standardized claim and its conformance assessment.

This example isolates the claim-model structure: stable identity bases, interaction-specific evidentiary binding, a typed claim drawn from the inherited taxonomy, and explicit claim boundaries that prevent an assertion from being interpreted beyond what was made.

3. Conformance — What It Means to Satisfy the Standard, and What Conformance Does Not Establish

Conformance determines whether an implementation satisfies the requirements defined by the standard. It is not itself an interpretation rule, nor is it a claim about the world.

A conformant implementation expresses claims according to the defined claim semantics, associates claims and evidence with specific interactions as required by interaction-specific evidentiary binding, declares claim boundaries as required, and satisfies the normative interpretation requirements. Conformance therefore establishes that an implementation meets the conditions under which cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries holds.

What conformance establishes is deliberately narrow. In particular, conformance does **not** establish:

- **That a claim is true.** A conformant implementation can emit a well-formed, correctly bound claim within its stated boundaries that is false. The standard does not determine whether a particular claim is true; conformance likewise does not transform satisfaction of the standard into a determination of veracity.
- **That the evidence is sufficient for a particular decision.** Sufficiency remains external to both the standardized claim and its conformance assessment (§4).
- **That conformant implementations will reach the same downstream inference or verdict.** Conformance establishes the conditions for agreement on the meaning, evidentiary binding, and boundaries of a claim. The Interaction Boundary Constraint does not prevent independently implemented systems from drawing different inferences about producing composition. Agreement on meaning is therefore not agreement on conclusion; only the former is established by the requirements defined here.
- **That an implementation's behavior is correct, safe, or aligned.** Those properties remain outside the conformance determination, consistent with the scope exclusions inherited from the Interaction Boundary Constraint and the Identity paper.

This distinction permits the standard to use *establish* without overclaiming. The object of agreement is deliberately constrained to structure **specified as part of the standardized claim**: what an assertion means, which interaction its evidence is bound to, and the boundaries attached to that assertion. Those are objects the standard can define and against which conformance can be determined. Downstream inference and evidentiary sufficiency are not. Extending the object of agreement to either would exceed what conformance can establish.

4. Reliance Boundary — Why Sufficiency for a Downstream Decision Remains External to the Standard and to Conformance

Sufficiency is relative in two dimensions: to the claim's stated boundaries and to the relying party's decision context.

The first dimension is **specified as part of the standardized claim**. It travels with the claim, the standard defines its meaning, and conformance can be determined against the requirements governing it.

The second dimension is **supplied externally at evaluation time** by the relying party. It is not part of the standardized claim and does not travel with it. This asymmetry—standardized claim content versus evaluation-time context—is the reason the line between this standard and downstream reliance falls where it does.

The distinction is checkable by inspection. Information specified as part of the standardized claim is governed by the standard. Information introduced externally to evaluate that claim against a particular decision context is not. Merely carrying information alongside a standardized claim does not make that information part of the standardized claim.

A determination that evidence is sufficient for a particular decision necessarily depends on this second dimension. The relying party may consider its risk tolerance, the stakes of the decision, the consequences of relying incorrectly, or other requirements particular to that decision context. Those considerations are not specified as part of the standardized claim. A standard that fixed sufficiency would therefore also have to fix some portion of the relying party's decision context, converting a cross-context interoperability standard into policy for a particular class of reliance decisions.

This separation has established analogues. The W3C Verifiable Credentials model (W3C, 2025) distinguishes verification of a credential from the verifier's subsequent application of its own policies and business rules before relying on the claims; a successfully verified credential need not be suitable for a particular use. The IETF Remote Attestation Procedures (RATS) architecture (Birkholz et al., 2023) makes the separation explicit through an Appraisal Policy for Attestation Results applied by the Relying Party when making application-specific decisions. In both cases, standardized information can be carried across implementations without standardizing the relying party's ultimate decision policy.

The resulting separation is the third layer of the paper's contract: **the standard defines the structure, meaning, and boundaries of the claim; conformance determines whether an implementation satisfies those requirements; reliance remains external to both.** This standard defines that seam and stops at it.

5. Operationalization — What an Implementation Must Emit, Preserve, or Expose So Another Implementation Can Agree on the Claim's Meaning, Binding, and Boundaries

The standard specifies what must be interpretable across independently implemented systems, not how an implementation is constructed. The requirements in this section define the observable surface of the interoperability contract. They do not prescribe an implementation architecture, evidence-production mechanism, or internal processing model.

The standardized information model describes the standardized claim, not the system that produces, stores, transports, or evaluates it. A standardized claim comprises, independent of representation: a claim typed according to the standardized claim semantics, its interaction-specific evidentiary binding, and its stated claim boundaries. The information model fixes the semantic meaning and relationships of those elements, not their serialization or conveyance. Normative interpretation requirements govern resolution of the model and are not themselves carried as content within the standardized claim.

The standard's semantic interoperability requirements apply once a concrete representation has been identified and mapped to those standardized information-model elements. Recognition of a concrete representation as an instance of the model, and mapping that representation to the model's elements, belong to representation profiles outside this standard. Any such mapping used to carry a standardized claim must preserve referential identity: representations of the same interaction referent must resolve to the same abstract interaction-referent value, and representations of distinct interaction referents must remain distinct. Given the same identified model elements, conformant implementations must resolve the same claim meaning, evidentiary binding, and stated claim boundaries.

Emit

When a conformant implementation asserts a standardized claim, it must make available, as part of that claim, the information required to interpret the assertion according to the normative objects defined in §2. This includes:

- the claim, typed according to the standardized claim semantics;
- its interaction-specific evidentiary binding; and
- its stated claim boundaries.

These elements must be explicitly represented rather than left for another implementation to infer. If information necessary to determine the claim's meaning, binding, or boundaries is absent from the standardized claim, cross-implementation agreement on that element cannot be established by the standard.

Preserve

The meaning and referential integrity of the standardized claim must remain stable when the claim is retained, transmitted, or interpreted across implementations. In particular, an implementation must preserve:

- resolution of component referents to the same stable identity bases inherited from the Identity paper;
- the identity of the declared interaction referent under representation-to-model mapping;
- the integrity of the evidentiary binding between the claim, its evidence, and the interaction it concerns; and
- the association between the assertion and its stated claim boundaries.

Preservation does not require a particular storage, transport, integrity, or identifier mechanism. It requires that those mechanisms, whatever their implementation, not alter the standardized meanings and relationships on which interpretation depends.

Expose

When a standardized claim is presented to a receiving implementation through a representation that has been mapped to the standardized information model, the implementation must be able to resolve the standardized information necessary to interpret the claim according to the normative interpretation requirements. It must be able to determine:

- which standardized claim type is asserted and the proposition that type asserts;
- which specific interaction the evidentiary binding resolves to;
- what stated claim boundaries constrain the assertion; and
- how the standard requires those elements to be interpreted.

Exposure therefore concerns interpretability, not internal visibility. A receiving implementation may internally partition or organize the underlying activity differently, but it must interpret the standardized claim at its declared interaction referent rather than substitute its own interaction unit. The standard does not require one implementation to reveal its architecture or internal state to another; it requires the standardized claim to expose enough information for an independently implemented receiver to resolve the same assertion, binding, and stated claim boundaries.

Emit, preserve, and expose describe complementary obligations of the interoperability contract. **Emit** requires the standardized objects to be explicitly represented. **Preserve** requires their meaning and relationships to remain invariant as the claim is retained, transmitted, or interpreted across implementations. **Expose** requires those objects to be resolvable by a receiving implementation. These obligations condition on the assertion and interpretation of standardized claims, not on permission to participate in an interaction. None determines whether the claim is true or whether its evidence is sufficient for a downstream decision.

Operational Contract Example

Standardized claim asserted → emit → preserve → expose → independent receiver resolves it

An orchestration service asserts a standardized observable-participation claim for interaction *K*. The claim states that an identified component, Model D, observably participated in *K*: an output, result, or action attributable to Model D entered the production of the declared interaction.

- **Emit** — The implementation makes available, as part of the standardized claim, the typed observable-participation assertion, its interaction-specific evidentiary binding to *K*, and its stated claim boundaries. The assertion is limited to Model D's observable participation in the identified interaction.
- **Preserve** — As the claim is retained or transmitted, Model D continues to resolve to the same identity basis, *K* continues to resolve to the same interaction-referent value, the evidentiary binding remains associated with *K*, and the relationship between the assertion and its stated claim boundaries remains unchanged. The mechanisms used to retain or transmit the claim do not alter these standardized relationships.
- **Expose** — When the standardized claim is presented to another independently implemented system, that system can resolve: (1) that an observable-participation claim is being asserted, (2) that the identified component is Model D, (3) that the evidentiary binding concerns interaction *K*, and (4) the stated claim boundaries under which the assertion is to be interpreted.

A receiving conformant implementation can therefore resolve the claim's meaning, its evidentiary binding, and its stated claim boundaries without needing visibility into the originating implementation's internal architecture or evidence-production process. It must interpret the claim at the declared interaction referent *K*, regardless of how it internally partitions the underlying activity, and it must not interpret observable participation as establishing material contribution, completeness, correctness, or another unstated property.

Whether the evidence is sufficient for a particular reliance decision remains external to the standardized claim and its conformance assessment.

This example isolates the operational surface of the interoperability contract: what must be emitted, what relationships must be preserved, and what must be exposed so that an independently implemented receiver can interpret the standardized claim according to the same claim semantics, evidentiary binding, and stated claim boundaries.

Prior-Art Positioning

The Interaction Claim Standard sits alongside established work on identity, verifiable claims, attestation, provenance, system documentation, and assurance arguments. These approaches provide mechanisms for identifying subjects, representing claims and evidence, documenting systems, and evaluating assurance. The standard neither

absorbs their machinery nor claims their territory. Its concern is cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries.

Decentralized Identifiers and Verifiable Credentials. The W3C Decentralized Identifiers (DIDs) standard (W3C, 2022) provides a common model for persistent identifiers and their resolution, while the Verifiable Credentials Data Model (W3C, 2025) provides a standardized structure for expressing verifiable claims about identified subjects. These mechanisms address stable identity and portable claims. They do not, by themselves, define what claims such as invocation, observable participation, completeness, or material contribution mean for a specific AI interaction, nor do they require supporting evidence to be bound to that interaction. The Interaction Claim Standard addresses a distinct question: standardized claim semantics, interaction-specific evidentiary binding, stated claim boundaries, and normative interpretation requirements for claims concerning a specific interaction.

Remote Attestation Procedures. The IETF Remote ATtestation procedureS (RATS) architecture (Birkholz et al., 2023) defines roles and information flows for remote attestation: an Attester produces Evidence about itself, a Verifier appraises that Evidence and produces an Attestation Result, and a Relying Party applies its own appraisal policy to that result. Evidence may describe configuration, measurements, telemetry, or other information about the Attester; the architecture organizes that Evidence around properties of the Attester or Target Environment rather than around participation in a specific interaction. The Interaction Claim Standard addresses the complementary problem: cross-implementation agreement on claims whose evidentiary binding resolves to a specific AI interaction.

RATS also provides precedent for the reliance boundary drawn here. The architecture separates the attestation information being appraised from the appraisal policies that determine how it is treated, and it makes that separation at both appraisal points: the Verifier applies its Appraisal Policy for Evidence, while the Relying Party applies its Appraisal Policy for Attestation Results. In each case, the policy is supplied by its respective owner rather than carried as part of the information being appraised. Evidence and Attestation Results can therefore travel across implementations without carrying the decision policy applied to them. The Interaction Claim Standard makes the analogous separation for interaction claims: the standardized claim carries its meaning, evidentiary binding, and stated claim boundaries; the decision context applied to that claim remains external.

Content provenance. The Coalition for Content Provenance and Authenticity (C2PA) Content Credentials specification (Coalition for Content Provenance and Authenticity, 2026) provides cryptographically verifiable provenance manifests associated with digital assets, recording assertions concerning an asset’s origin and history. Its evidentiary association is therefore principally artifact-bound: the manifest describes the provenance of the asset to which it is associated. The Interaction Claim Standard addresses a different binding problem. Its standardized claims concern a specific AI interaction, and the evidence supporting those claims must resolve to that interaction rather than merely to an artifact, component, configuration, or provenance record

associated with it. Artifact provenance can therefore provide relevant information without, by itself, establishing an interaction-specific claim.

Model Cards. Model Cards (Mitchell et al., 2019) provide structured documentation describing a model's intended use, evaluation conditions, performance characteristics, and known limitations. Such documentation supplies information about a model independently of any particular interaction in which that model may participate. The Interaction Claim Standard addresses the interaction-indexed case: what claim is being asserted about a specific interaction, which interaction the supporting evidence concerns, and the stated claim boundaries of that assertion. System- or model-level documentation can inform interpretation, but it does not substitute for evidence bound to the interaction under evaluation.

Assurance-case notations. Claims-Arguments-Evidence (CAE) (Bloomfield & Netkachova, 2014) and Goal Structuring Notation (GSN) (SCSC Assurance Case Working Group, 2021) provide structured representations for arguments connecting claims and evidence: GSN standardizes a notation for developing and evaluating engineering arguments, including the relationships among claims, evidence, and context, and CAE structures assurance arguments around its three named elements. The Interaction Claim Standard deliberately standardizes no such argument. It fixes the meaning, evidentiary binding, and stated claim boundaries of assertions as part of the requirements for cross-implementation agreement on the meaning of claims, their evidentiary bindings, and their stated boundaries. Any argument from a standardized claim and its evidence to a downstream conclusion remains unstandardized and, where constructed, belongs to the relying party's decision context. The absence of a standardized evidentiary argument is a deliberate exclusion from the standard, not an unaddressed component.

References

Birkholz, H., Thaler, D., Richardson, M., Smith, N., & Pan, W. (2023). *Remote ATtestation procedureS (RATS) Architecture* (RFC 9334). Internet Engineering Task Force. https://www.rfc-editor.org/rfc/rfc9334

Bloomfield, R., & Netkachova, K. (2014). Building blocks for assurance cases. In *2014 IEEE International Symposium on Software Reliability Engineering Workshops* (pp. 186–191). IEEE. https://doi.org/10.1109/ISSREW.2014.72

Coalition for Content Provenance and Authenticity. (2026). *C2PA Technical Specification* (Version 2.4, April 2026). https://spec.c2pa.org/specifications/specifications/2.4/specs/C2PA_Specification.html

Harber, B. (2026a). *Interaction Boundary Constraint.* Technical Disclosure Commons, Defensive Publications Series. https://www.tdcommons.org/dpubs_series/10836/

Harber, B. (2026b). *Toward a Standard of Identity in AI Interaction.* Technical Disclosure Commons, Defensive Publications Series. https://www.tdcommons.org/dpubs_series/11309/

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287596

SCSC Assurance Case Working Group. (2021). *Goal Structuring Notation Community Standard Version 3* (SCSC-141C). Safety-Critical Systems Club. https://scsc.uk/gsn

World Wide Web Consortium. (2022). *Decentralized Identifiers (DIDs) v1.0.* W3C Recommendation, 19 July 2022. https://www.w3.org/TR/did-core/

World Wide Web Consortium. (2025). *Verifiable Credentials Data Model v2.0.* W3C Recommendation, 15 May 2025. https://www.w3.org/TR/vc-data-model-2.0/