Encevra Research Round 36: When One Complete Document Is Still Half the Truth: Cross-Document Reference Closure, Precedence DAGs, and Compositional Clause Chains Across Multi-Instrument Quorums

Date: 2026-10-28
Research lab: Encevra Research
Primary tracks: Track 1 (Evidence, Provenance, and Change), Track 2 (Selective Verification Under Budget), Track 5 (Calibration, Abstention, and Release Control)
Carry-forward boundary: Round 35 defined Correlated-Inspection Sufficiency (`CIS-35`), pairing a `4,096`-state deterministic post-`GTP-34` diagnostic (`E35-A`) with a `100`-evaluation live multi-model contract-audit benchmark (`E35-B`) across four distinct witness architecture families (`Witness A` on Max reasoning, `Witness B`, `Witness C`, and `Witness D`). While `CIS-35` eliminated shared retrieval-slice false consensus within a single canonical contract bundle (`0 / 12` false certified releases), Round 35 Section 9.1 (`F35.5`) and Section 11 explicitly handed off two unresolved boundaries: (1) `CIS-35` suffered `2 / 12` (`16.67%`) conservative fail-closed holds (`C03` and `C12`) because its single scalar `governing_section` schema could not represent a two-clause composition (`base_entitlement_section` plus `decisive_override_section`) or a signed inline fact attestation (`inline_fact_sufficiency`), and (2) `CIS-35` verifies byte completeness of a single primary document (`h_i = SHA256(D_primary)`), leaving open what happens in enterprise transactions governed by a multi-document reference graph (`G_doc = (D, E_ref, pi_prec)`) where a 100% locally complete Master Services Agreement or Data Processing Addendum incorporates an unattached Order Form, Rider, Schedule, Side Letter, or Amendment that inverts the governing obligation.

Abstract

In enterprise legal, regulatory, and infrastructure assurance workflows, obligations rarely reside inside a single self-contained file. Instead, an obligation is governed by a directed reference graph `G_doc = (D, E_ref, pi_prec)` spanning a base Master Services Agreement (`DOC-MSA`), Data Processing Addendum (`DOC-DPA`), Service Level Agreement (`DOC-SLA`), executed Order Forms (`DOC-OF`), Schedules, Closing Side Letters (`DOC-CSL`), and chronological Amendments (`DOC-AMD`), resolved under a contractual order-of-precedence hierarchy `pi_prec`. Round 35 proved that within a single canonical bundle, requiring full-bundle digest binding and verbatim governing-span extraction (`CIS-35`) eliminates retrieval-slice truncation errors (`0 / 12` false releases), but left a `16.67%` (`2 / 12`) false-hold rate from rigid scalar clause labeling (`C03`, `C12`) and did not test cross-document incorporation by reference where a witness panel receives a byte-complete primary agreement (`V_scope_local = 1`) whose companion overriding instrument is unattached or superseded.

Round 36 introduces **Cross-Document Completeness and Reference-Graph Sufficiency (`CCS-36`)** and evaluates it through two strictly separated evidence categories: an exact `16,384`-state deterministic structural and combinatorial diagnostic (`E36-A`) and a `108`-evaluation live matched-compute multi-model benchmark (`E36-B`) across `12` multi-instrument enterprise transaction packets (`M01-M12`) using four distinct witness architecture families (`Witness A` on Max reasoning `reasoning: {"effort": "high"}`, `Witness B`, `Witness C`, and `Witness D`).

First, in the deterministic diagnostic (`E36-A`, `4^7 = 16,384` multi-document reference-graph and witness-quorum configurations), we prove two structural results: (i) **Theorem 36.1 (Local-Bundle Completeness Insufficiency Under Cross-Document Override)**, showing that applying Round 35's `CIS-35` gate to a locally hash-complete primary document (`G1_cis35_local_single_doc`) still permits `126` false certifications (`FAR = 0.7694%`) whenever an unattached or precedence-inverted companion instrument reverses the base clause, while simultaneously rejecting `6 / 8` (`FRR = 75.00%`) valid multi-document compositions due to rigid scalar clause labeling and inline-exhibit rejection; and (ii) **Theorem 36.2 (`CCS-36` Reference-Closure and Precedence-DAG Sufficiency)**, showing that the conjunction of transitive reference-closure verification (`V_ref_closure`), acyclic order-of-precedence DAG resolution (`V_prec_dag`), compositional clause-chain tuple verification (`V_tuple`), cross-document verbatim span grounding (`V_xspan`), active temporal execution status (`V_temp`), and `⩾ 3` witness architecture families (`G5_ccs36_full`) achieves `TP = 8`, `FP = 0`, `TN = 16,376`, `FN = 0` (`0.00%` false acceptance rate, `0.00%` false rejection rate, `100.00%` accuracy). An exact multi-hop reference-truncation derivation further demonstrates that at marginal independent witness error `p_indep = 0.15` (`1.20%` independent 3-of-4 quorum tail), a per-hop cross-document retrieval recall of `q = 0.85` across a 3-document chain (`m = 3`) induces a `27.75%` common-mode truncation shock and raises 3-of-4 quorum failure to `28.62%`, requiring per-hop recall `q > q* = 96.50%` (`0.964964`) to outperform a single full-packet auditor (`p_single = 0.08`).

Second, in the live multi-model benchmark (`E36-B`, `108` live witness evaluations across `12` multi-document packets `M01-M12`), we evaluate five arms across `Witness A` (`54` calls), `Witness B` (`18` calls), `Witness C` (`18` calls), and `Witness D` (`18` calls). On Stratum 1 (`M01-M06`, Cross-Document Dangling Override Traps where a 100% locally complete primary agreement incorporates an unattached Order Form, Sovereign Banking Rider, High-Availability Addendum, Cloud Hosting Schedule, Closing Side Letter, or Commercial Reset Amendment), both local single-document `CIS-35` (`Arm 2`) and an ungated 4-family heterogeneous quorum (`Arm 3`) suffer `6 / 6` (`100.00%`) unanimous `4-0` false certified releases (`6 / 12 = 50.00%` overall false certification rate). Furthermore, on Stratum 2 (`M07-M12`), local `CIS-35` (`Arm 2`) again holds `2 / 12` (`16.67%`) valid packets fail-closed on the `C03` and `C12` compositional/inline-attestation successors (`M09` and `M10`), leaving `Arm 2` with only `4 / 12` (`33.33%`) correct certified releases. Enforcing `CCS-36` before cross-document bundle reopen (`Arm 4`) intercepts `6 / 6` dangling-reference packets fail-closed (`0 / 12` false releases) because all `4 / 4` witness families flag `unattached_external_reference = true` and `V_ref_closure` detects the missing target digests. Once `CCS-36` closes the transitive multi-document bundle (`Arm 5`), all `6 / 6` Stratum-1 packets flip from `4-0` false consensus to `4-0` verified correct consensus, and both `M09` and `M10` resolve cleanly without false holds, achieving `12 / 12` (`100.00%`) certified correct releases, `0 / 12` (`0.00%`) false certified releases, and `0 / 12` (`0.00%`) conservative holds.

---

1. Archive continuity and anti-repetition verification across Rounds 1-35

Before formulating Round 36 or executing any diagnostic or live witness call, we reviewed all `35` prior rounds in the Encevra SI Research Labs chronology (`29` indexed archive rounds through Round 35).

Round	Date	Core Question and Mechanism	Explicit Carry-Forward Boundary to Round 36
Round 31	2026-10-22	Can a signed sandbox prove single file-backed root interventions before model inference (`RISE`)?	Proved single-action file fidelity (`7/7` conformance); left multi-service collusion open.
Round 32	2026-10-23	Can separately keyed witness, registry, and monitor services catch blinded mutations?	Caught `208/224` single-boundary faults, `0/16` shared-admin aliases (`0` live model calls).
Round 33	2026-10-24	What provenance completeness is needed before a 3-of-4 witness quorum is certified (`AIC-33`)?	`3,375` synthetic states (`0` live model calls); noted structural independence does not guarantee uncorrelated model errors.
Round 34	2026-10-25	When cross-certificates chain, what separates grounded trust paths from circular/Sybil loops (`GTP-34`)?	`279,936` synthetic graphs (`0` live model calls); handed off correlated semantic failures among grounded witnesses.
Round 35	2026-10-27	When a 3-of-4 quorum satisfies `AIC-33` and `GTP-34`, when do live model witnesses still converge on a false verdict (`CIS-35`)?	`4,096` post-`GTP-34` states (`E35-A`) + `100` live evaluations (`E35-B`); eliminated intra-bundle slice truncation (`0/12` false releases) but left `2/12` (`16.67%`) scalar/inline-exhibit holds (`C03`, `C12`) and did not test cross-document reference graphs where a locally complete base agreement is overridden by an unattached companion instrument.
Round 36	2026-10-28	When each witness inspects a 100% byte-complete primary agreement inside a multi-document transaction graph `G_doc = (D, E_ref, pi_prec)`, how do dangling incorporation-by-reference edges, precedence inversions, and compositional clause chains cause false certifications or false holds, and does `CCS-36` achieve `0/12` false releases and `0/12` false holds at matched compute?	Combines an exact `16,384`-state multi-document reference-graph diagnostic (`E36-A`) with a `108`-evaluation live 4-architecture multi-document benchmark (`E36-B`).

2. Executive summary

1. **Local single-document completeness (`V_scope_local = 1`) does not imply cross-document decision sufficiency (`V_ref_closure = 1`).** Under Round 35's `CIS-35` gate, if an assurance pipeline verifies that all four witnesses inspected 100% of the bytes of a Master Services Agreement (`h_i == SHA256(DOC-MSA)`), the panel passes local digest verification. However, if `DOC-MSA Section 8.1` states *"Except as expressly superseded in an applicable Order Form..."* and `Order Form #4` (`DOC-OF4`) is stored as a separate instrument in the closing binder and omitted from the witness packet, all four witness architecture families faithfully apply the base MSA cap and converge `4-0` on an inverted audit verdict.
2. **Rigid scalar clause attribution (`governing_section`) causes avoidable fail-closed holds on valid multi-clause compositions.** Round 35's `E35-B` benchmark recorded a `16.67%` (`2 / 12`) conservative hold rate (`C03` and `C12`) because witnesses were forced to emit a single scalar section label when the legal conclusion required composing a base entitlement clause (`DOC-SLA Section 1.2`) with a quantitative exclusion cap in an exhibit (`DOC-EXB Section 3.4`), or verifying a cryptographically signed inline fact attestation (`INLINE-ATTESTED-EX-C`).
3. **Evidence Category 1 - Exact `16,384`-state multi-document structural diagnostic (`E36-A`):** Across `4^7 = 16,384` post-`GTP-34` multi-document configurations (`8` positive `CCS_CERTIFIED` states and `16,376` negative compromised states), releasing on `GTP-34` 3-of-4 verdict agreement alone (`G0`) yields `4,088` false certifications (`FAR = 24.96%`). Applying local single-document `CIS-35` (`G1_cis35_local_single_doc`) still permits `126` false certifications (`FAR = 0.77%`) on cross-document overrides while rejecting `6 / 8` (`FRR = 75.00%`) valid multi-document compositions. Multi-document attachment without precedence-DAG verification (`G2`) leaves `56` false certifications, and rigid scalar schemas on multi-document bundles (`G3`) eliminate false certifications (`FP = 0`) only at the cost of a `75.00%` (`6 / 8`) false rejection rate. Only the full **`CCS-36`** gate (`G5_ccs36_full`) achieves `TP = 8`, `FP = 0`, `TN = 16,376`, `FN = 0` (`0.00%` FAR, `0.00%` FRR, `100.00%` accuracy).
4. **Evidence Category 1 - Multi-hop cross-document truncation surface (`E36-A`):** When an obligation spans `m` cross-referenced instruments (`m - 1` reference hops) and an upstream document-graph retriever resolves each hop with probability `q in (0, 1]`, the probability of transitive bundle closure is `P_closed(m, q) = q^(m - 1)`, inducing a common-mode truncation shock `rho_trunc(m, q) = 1 - q^(m - 1)`. At independent witness error `p_indep = 0.15` (`1.20%` independent 3-of-4 quorum failure), a seemingly high per-hop recall of `q = 0.85` across `m = 3` documents causes 3-of-4 quorum failure to jump `23.9x` from `1.20%` to `28.62%` (and to `48.42%` at `m = 5`). Across `m = 3` documents, a 3-of-4 quorum is strictly less reliable than a single full-packet auditor (`p_single = 0.08`) whenever per-hop retrieval recall falls below `q* = 96.50%` (`0.964964`).
5. **Evidence Category 2 - Live 4-architecture multi-document benchmark (`E36-B`):** Across `12` multi-document enterprise transaction packets (`M01-M12`) and `108` live witness evaluations (`Witness A` on Max reasoning `effort: high` with `54` calls, `Witness B` with `18` calls, `Witness C` with `18` calls, and `Witness D` with `18` calls), both local single-document `CIS-35` (`Arm 2`) and an ungated 4-family quorum (`Arm 3`) suffer `6 / 6` (`100.00%`) unanimous `4-0` false certified releases on Stratum 1 (`M01-M06`, cross-document dangling overrides), yielding a `50.00%` (`6 / 12`) overall false certification rate. In addition, `Arm 2` holds `M09` and `M10` fail-closed (`2 / 12 = 16.67%` hold rate), certifying only `4 / 12` (`33.33%`) packets correctly.
6. **Live empirical validation of `CCS-36` (`Arm 4` and `Arm 5`):** On the partial packets (`Arm 4`), `CCS-36` catches all `6 / 6` dangling cross-document overrides fail-closed (`HOLD_CCS36_DANGLING_REF_CLOSURE`, `0 / 12` false certified releases) and already certifies all `6 / 6` complete Stratum-2 packets (`M07-M12`, including `M09` and `M10`). Once `CCS-36` triggers transitive reference-graph closure to attach the missing companion instruments (`Arm 5`), all `12 / 12` (`100.00%`) packets achieve `4-0` cross-architecture consensus with verified compositional clause tuples and verbatim spans (`12 / 12` correct certified releases, `0 / 12` false certified releases, `0 / 12` fail-closed holds).

3. Problem statement

Consider an enterprise assurance controller auditing a transaction question `q` over a contractual corpus `D = {D_1, D_2, ..., D_m}`. Suppose the witness panel `W = {W0, W1, W2, W3}` already satisfies Round 33 (`AIC-33` authority independence), Round 34 (`GTP-34` grounded trust paths with `|R_G(Q)| ≥ 3`), and Round 35's local document digest check on the primary instrument `D_1` (`h_i == SHA256(D_1)`).

Why can a 3-of-4 heterogeneous witness panel still fail systematically on multi-instrument enterprise packets? Three structural mechanisms govern this boundary:

1. **Cross-Document Incorporation-by-Reference Truncation (`E_ref` dangling edges):** A primary instrument `D_1` (for example, `DOC-MSA` or `DOC-DPA`) is attached in full (`100%` of `D_1`'s bytes are present, so no intra-document chunk truncation occurred). However, `D_1` contains an explicit or conditional reference edge `(D_1, D_k) in E_ref` pointing to a separate legal instrument `D_k` (such as `Order Form #4`, `Sovereign Banking Rider`, `High-Availability Addendum`, `Schedule C`, `Closing Side Letter`, or `Second Amendment`). If `D_k` is absent from the inspected packet `B_attached subset D`, instruction-following witnesses cannot fabricate `D_k`'s text; they apply the default rule in `D_1` and unanimously emit the inverted verdict (`4-0`).
2. **Order-of-Precedence DAG Inversions and Specificity Guardrails (`pi_prec`):** When multiple instruments `{D_1, D_2, D_3}` are attached simultaneously, they frequently contain facially conflicting terms: for example, an Order Form (`DOC-OF2`) sets a general commercial cap of `500,000 units` while a Data Processing Addendum (`DOC-DPA`) sets a `5,000,000 unit` super-cap for privacy claims under a precedence hierarchy `DOC-DPA > DOC-OF2 > DOC-MSA` (`M07`), or a Statement of Work (`DOC-SOW3`) purports to assign all IP to Supplier without citing `DOC-MSA Section 6.1` as required by the MSA's Precedence Specificity Rule (`M08`), or an unexecuted draft side letter (`DOC-DRAFT-SL`) or superseded First Amendment (`DOC-AMD1`) sits in the closing binder alongside the active executed instrument (`M11`, `M12`).
3. **Scalar vs. Compositional Clause-Chain Schemas (`C03` and `C12` bottleneck from Round 35):** When a decision requires combining a base entitlement rule in `D_1` (`DOC-SLA Section 1.2`) with a quantitative cap or exception in `D_2` (`DOC-EXB Section 3.4`), witnesses that agree `4-0` on the reasoning chain split on whether to name `D_1` or `D_2` in a single scalar `governing_section` slot (`M09`). Similarly, when an external list (`Exhibit C`) is stipulated via a cryptographically signed inline registry attestation (`INLINE-ATTESTED-EX-C`), a binary `missing_cross_reference` flag without an explicit `inline_fact_sufficiency` predicate triggers unnecessary fail-closed holds (`M10`).

---

4. Literature review and theoretical grounding (`[S01]`–`[S12]`)

Round 36 builds on the fault-tolerant quorum and multi-agent reliability literature established in Round 35 while adding primary sources on multi-hop retrieval completeness, relational document graphs, and compositional verification:

- **Dependent failures and common-mode input shocks (`[S01]`–`[S05]`):** Eckhardt and Lee (`[S02]`, 1985), Knight and Leveson (`[S01]`, 1986), and Littlewood and Miller (`[S03]`, 1995) established that architectural diversity across checkers does not bound majority-vote error when checkers share a common failure-inducing input condition. Malkhi and Reiter (`[S04]`, 1998) and Junqueira and Marzullo (`[S05]`, 2003) proved that quorum safety depends on bounding correlated failure sets rather than counting votes.
- **Multi-agent LLM consensus and verification limits (`[S06]`–`[S12]`):** While self-consistency (`[S06]`, Wang et al., 2023) and multi-agent debate (`[S07]`, Du et al., 2023; `[S08]`, Liang et al., 2023; `[S10]`, Li et al., 2024) improve performance on self-contained tasks, Huang et al. (`[S09]`, 2023) showed that language models cannot reliably self-correct without grounded external verifier signals, Wu et al. (`[S11]`, 2024) quantified shared model blind spots, and Cemri et al. (`[S12]`, 2025) demonstrated that multi-agent LLM panels systematically compound errors when upstream context or specification handoffs are incomplete.

---

5. Formal model and theorems: Cross-Document Completeness & Sufficiency (`CCS-36`)

5.1 Multi-document reference graph and precedence DAG

Let an enterprise transaction packet be represented as a directed multi-document graph:

```
G_doc = (D, E_ref, pi_prec, tau_exec)
```

where: - `D = {D_1, ..., D_m}` is the finite set of candidate instruments in the transaction binder, each with canonical digest `H(D_r) = SHA256(D_r)`. - `E_ref subset D x D` is the set of directed incorporation-by-reference edges: `(D_a, D_b) in E_ref` means a clause in `D_a` expressly conditions, delegates, or subordinates its legal effect to instrument `D_b` (or to an inline-attested registry fact `A_inline(D_b)`). For a query `q` rooted at base instrument `D_1`, let `Cl(q, G_doc) subset D` denote the transitive reference closure of instruments reachable from the clauses relevant to `q`. - `pi_prec` is a partial order (directed acyclic graph) over `D` specifying which instrument prevails in the event of a conflict on subject-matter domain `q`, optionally guarded by a specificity predicate `Spec(D_high, D_low, sec_id)` requiring `D_high` to cite `sec_id` of `D_low` explicitly (as in `M08`). - `tau_exec : D → {ACTIVE_EXECUTED, SUPERSEDED, UNEXECUTED_DRAFT}` labels the execution and temporal validity status of each instrument in `D`.

5.2 Multi-hop reference-truncation shock and Theorem 36.1

Suppose a query `q` depends on a chain of `m ⩾ 1` documents `Cl(q, G_doc) = {D_1, D_2, ..., D_m}` linked by `m - 1` cross-document incorporation hops, where `D_m` contains a decisive override that reverses the default verdict of `D_1`. If an upstream document assembler includes `D_1` with probability `1` (`V_scope_local(D_1) = 1`) and independently resolves each cross-document reference hop `(D_r, D_{r+1})` with per-hop recall `q in (0, 1]`, then the probability that the attached packet `B_attached` contains the complete transitive closure `Cl(q, G_doc)` is:

```
P_closed(m, q) = q^(m - 1)
```

Whenever any hop fails (`probability rho_trunc(m, q) = 1 - q^(m - 1)`), the overriding instrument `D_m` is absent from `B_attached`, inducing a `100%` common-mode premise shock across all instruction-following witnesses inspecting `B_attached`. For a 3-of-4 witness panel whose conditional error rate on a complete packet is `p_indep in (0, 1)` (with independent 3-of-4 quorum failure tail `P_tail(p_indep) = 4 * p_indep^3 * (1 - p_indep) + p_indep^4`), the ungated 3-of-4 quorum failure probability is:

```
P_fail_quorum(m, q, p_indep) = (1 - q^(m - 1)) + q^(m - 1) * (4 * p_indep^3 * (1 - p_indep) + p_indep^4)
```

Theorem 36.1 (Local-Bundle Completeness Insufficiency Under Cross-Document Override). *Let `W = {W0, W1, W2, W3}` be a witness panel satisfying `AIC-33`, `GTP-34` (`|R_G(Q)| ≥ 3`), and `≥ 3` distinct witness architecture families. If the release verifier checks only local primary-document completeness (`h_i == SHA256(D_1)`) and local verbatim span grounding (`q_i substring D_1`) without enforcing transitive cross-document reference closure (`Cl(q, G_doc) subset B_attached`), then on every transaction where an unattached higher-precedence instrument `D_k in Cl(q, G_doc) \ B_attached` reverses `D_1`, faithful instruction-following witnesses emit unanimous `4-0` false consensus, and the critical per-hop retrieval recall `q*` required for an ungated 3-of-4 quorum across `m` documents to match a single full-packet auditor with error `p_single > P_tail(p_indep)` is:*

```
q*(m, p_indep, p_single) = ((1 - p_single) / (1 - P_tail(p_indep)))^(1 / (m - 1))
```

Proof sketch. When `D_k \notin B_attached`, the visible text in `D_1` states a base rule `v_base` subject to `D_k`. Because witnesses are prohibited from hallucinating unattached text, every faithful witness evaluates `D_1` to `v_base ≠ v*`, yielding `4-0` agreement on `v_base` with a valid verbatim quote from `D_1` and a valid local hash `SHA256(D_1)`. Setting `P_fail_quorum(m, q*, p_indep) = p_single` and solving `1 - q*^(m - 1) * (1 - P_tail(p_indep)) = p_single` yields the closed-form threshold `q*(m, p_indep, p_single)` above. For `m = 3`, `p_indep = 0.15` (`P_tail = 0.01198125`), and `p_single = 0.08`, `q* = (0.92 / 0.98801875)^(1/2) = 0.964964` (`96.50%` per-hop recall).

5.3 The `CCS-36` compositional receipt schema and Theorem 36.2

To eliminate both cross-document dangling overrides (`M01-M06`) and scalar clause-label false holds (`M09`, `M10`), `CCS-36` requires each witness `W_i` to emit a compositional cross-document inspection receipt:

```
r_i = (v_i, s_base_i, s_ctrl_i, pi_i, q_i, u_i, a_inline_i, M_i)
```

where: - `v_i in {PASS, FAIL}` is the binary audit verdict, - `s_base_i = (doc_base_i, sec_base_i)` identifies the base entitlement or default clause in the primary instrument, - `s_ctrl_i = (doc_ctrl_i, sec_ctrl_i)` identifies the decisive controlling or overriding clause after cross-document precedence resolution (equal to `s_base_i` when no override applies), - `pi_i` is the witness's verified precedence relation (`doc_ctrl_i > doc_base_i`), - `q_i` is a contiguous verbatim substring (`|norm(q_i)| ≥ 18`) extracted from the controlling instrument, - `u_i in {true, false}` (`unattached_external_reference`) flags whether any external instrument incorporated by reference in the visible text is missing from the attached bundle and not covered by a signed inline fact attestation, - `a_inline_i in {true, false}` (`inline_fact_sufficiency`) records whether an external schedule reference is satisfied by a cryptographically signed inline registry fact attestation, and - `M_i = {SHA256(D) : D in B_attached}` is the manifest of cryptographic digests of all attached instruments inspected by `W_i`.

Theorem 36.2 (Reference-Closure and Compositional Precedence-DAG Sufficiency `CCS-36`). *A 3-of-4 agreeing witness subset `Q subset W (|Q| ≥ 3)` is certified for release (`CCS_CERTIFIED`) if and only if all six predicates hold:*

1. **Grounded Multi-Family Independence (`V_arch`):** `Q` satisfies `AIC-33` and `GTP-34` (`|R_G(Q)| ≥ 3`) and spans at least three distinct witness architecture families (`|{family(W_i) : W_i in Q}| ≥ 3`).
2. **Transitive Reference-Graph Closure (`V_ref_closure`):** Every required instrument `D_r in Cl(q, G_doc)` is either cryptographically present in the attached manifest `M_i` (`SHA256(D_r) in M_i`) or satisfied by a signed inline fact attestation (`a_inline_i == true` verified against the registry receipt), and fewer than two witnesses in `W` flag `u_j == true`.
3. **Acyclic Order-of-Precedence DAG Verification (`V_prec_dag`):** The precedence relation `pi_i` cited by witnesses in `Q` is acyclic and consistent with the governing hierarchy `pi_prec` and any specificity guardrail `Spec(doc_ctrl, doc_base, sec_base)`.
4. **Active Temporal Execution Verification (`V_temp`):** The controlling instrument `doc_ctrl_i` has status `tau_exec(doc_ctrl_i) == ACTIVE_EXECUTED` (rejecting `SUPERSEDED` prior amendments and `UNEXECUTED_DRAFT` side letters).
5. **Compositional Clause-Chain Tuple Reconciliation (`V_tuple`):** At least three witnesses in `Q` match the canonical compositional chain `(s_base*, s_ctrl*)`, allowing witnesses to record both the base entitlement clause (`s_base_i`) and the decisive ceiling/override clause (`s_ctrl_i`) without scalar-slot collision.
6. **Cross-Document Verbatim Span Verification (`V_xspan`):** At least three witnesses in `Q` provide a contiguous verbatim quote `norm(q_i)` verified against the active controlling instrument in the complete multi-document bundle.

When any predicate fails because `V_ref_closure == 0`, `CCS-36` transitions fail-closed to `HOLD_CCS36_DANGLING_REF_CLOSURE` and triggers a targeted bundle reopen over `Cl(q, G_doc)`. In the finite `16,384`-state universe of `E36-A`, the conjunction `V_arch & V_ref_closure & V_prec_dag & V_temp & V_tuple & V_xspan` is both necessary and sufficient for `0` false acceptances (`FP = 0 / 16,376`) and `0` false rejections (`FN = 0 / 8`).

---

6. Hypotheses (`H36.1` - `H36.7`)

- **`H36.1` (Local-Bundle `CIS-35` Insufficiency Under Cross-Document Override):** Applying Round 35's `CIS-35` gate to a 100% locally complete primary document without cross-document reference-graph closure (`G1_cis35_local_single_doc` in `E36-A; Arm 2` in `E36-B`) permits unanimous false certifications whenever an unattached higher-precedence companion instrument reverses the base obligation (`126` false acceptances in `E36-A`; `6 / 6` false certified releases on Stratum 1 `M01-M06` in `E36-B`).
- **`H36.2` (Multi-Hop Reference-Truncation Amplification):** Across an `m`-document reference chain with per-hop retrieval recall `q = 0.85` and independent witness error `p_indep = 0.15`, compound reference-hop losses elevate ungated 3-of-4 quorum failure from `1.20%` (`m = 1`) to `16.02%` (`m = 2`), `28.62%` (`m = 3`), `39.32%` (`m = 4`), and `48.42%` (`m = 5`), requiring `q > q* = 96.50%` at `m = 3` to beat a single `p_single = 0.08` full-packet checker.
- **`H36.3` (Multi-Document Attachment Insufficiency Without Precedence-DAG and Temporal Verification):** Attaching all referenced documents and checking verbatim spans without verifying the contractual Order-of-Precedence DAG and active amendment execution status (`G2_multidoc_attachment_no_prec_dag` in `E36-A`) still falsely accepts `56` compromised states where witnesses apply a subordinate, superseded, or unexecuted clause.

- `H36.4` **(Scalar Schema False-Hold Bottleneck vs. Compositional Tuple Resolution):** Enforcing a rigid single scalar `governing_section` field on multi-document bundles (`G3_multidoc_scalar_schema_rigid` in `E36-A`; `Arm 2` on `M09` and `M10` in `E36-B`) causes a `75.00%` (`6 / 8`) false rejection rate in `E36-A` and `2 / 12` (`16.67%`) false holds in `E36-B`, whereas `CCS-36`'s compositional `(base_entitlement_section, decisive_override_section, inline_fact_sufficiency)` tuple reduces false holds to `0 / 8` in `E36-A` and `0 / 12` in `E36-B`.
- `H36.5` **(Deterministic Exactness of Full `CCS-36`):** Across all `4^7 = 16,384` enumerated multi-document states in `E36-A`, the complete `CCS-36` gate (`G5_ccs36_full`) achieves `TP = 8`, `FP = 0`, `TN = 16,376`, `FN = 0` (`0.00%` FAR, `0.00%` FRR, `100.00%` accuracy).
- `H36.6` **(Live Fail-Closed Interception of Dangling Cross-Document References):** On the 12-packet live 4-architecture benchmark (`E36-B`), applying `CCS-36` prior to bundle reopen (`Arm 4`) intercepts `6 / 6` Stratum-1 dangling-override packets (`M01-M06`) fail-closed (`HOLD_CCS36_DANGLING_REF_CLOSURE`) with `0 / 12` false certified releases, while ungated 4-family voting (`Arm 3`) and local `CIS-35` (`Arm 2`) both suffer `6 / 6` unanimous `4-0` false certified releases on `M01-M06`.
- `H36.7` **(Live 12/12 Recovery Under Transitive Bundle Closure and Compositional Schemas):** Once `CCS-36` reopens the transitive multi-document bundle (`Arm 5` in `E36-B`), all `6 / 6` Stratum-1 packets flip from `4-0` false consensus to `4-0` verified correct consensus and all `6 / 6` Stratum-2 precedence-DAG and compositional packets (`M07-M12`, including the `C03` and `C12` successors `M09` and `M10`) release cleanly, achieving `12 / 12` (`100.00%`) correct certified releases, `0 / 12` (`0.00%`) false certified releases, and `0 / 12` (`0.00%`) conservative holds.

---

7. Evidence Category 1: Deterministic structural diagnostic (`E36-A`)

7.1 Experimental design (`4^7 = 16,384` multi-document reference-graph configurations)

Kept strictly separate from live model evaluations, the deterministic diagnostic `E36-A` enumerates the cartesian product of seven 4-state dimensions (`4^7 = 16,384` total states):

1. `local_doc_integrity`: {`0_truncated_base_doc`, `1_unverified_summary_only`, `2_locally_complete_base_doc_hash_verified`, `3_full_multi_doc_manifest_hash_verified`}.
2. `reference_closure_state`: {`0_dangling_silent_omission`, `1_dangling_flagged_unreopened`, `2_inline_fact_attested_closure`, `3_full_transitive_bundle_closure`}.
3. `precedence_dag_state`: {`0_precedence_ignored_base_wins`, `1_circular_or_ambiguous_priority`, `2_lexical_recency_heuristic_only`, `3_verified_acyclic_precedence_dag`}.
4. `clause_chain_schema`: {`0_ungrounded_verdict_only`, `1_scalar_base_section_only`, `2_scalar_label_split_on_valid_chain`, `3_verified_compositional_tuple`}.
5. `cross_doc_span_grounding`: {`0_paraphrased_no_verbatim_match`, `1_base_doc_span_verified_override_unverified`, `2_decoy_or_superseded_redline_span`, `3_both_base_and_controlling_spans_verified_active`}.
6. `architecture_diversity`: {`0_single_witness_no_quorum`, `1_homogeneous_monoculture_4w`, `2_dual_family_quorum`, `3_quad_family_grounded_quorum`}.
7. `temporal_amendment_status`: {`0_superseded_prior_version_applied`, `1_unexecuted_draft_side_letter_applied`, `2_active_initial_instrument_no_amendment`, `3_active_latest_executed_amendment_verified`}.

A state is ground-truth `CCS_CERTIFIED` positive (`8` states total) if and only if: - `local_doc_integrity == 3` (`full_multi_doc_manifest_hash_verified`), - `reference_closure_state in {2, 3}` (`inline_fact_attested_closure` or `full_transitive_bundle_closure`), - `precedence_dag_state == 3` (`verified_acyclic_precedence_dag`), - `clause_chain_schema in {2, 3}` (valid two-clause chain, whether witnesses split across the two labels under a scalar schema or emit the unified compositional tuple), - `cross_doc_span_grounding == 3` (`both_base_and_controlling_spans_verified_active`), - `architecture_diversity == 3` (`quad_family_grounded_quorum`), and - `temporal_amendment_status in {2, 3}` (`active_initial_instrument_no_amendment` or `active_latest_executed_amendment_verified`).

All remaining `16,384 - 8 = 16,376` states contain at least one cross-document truncation, precedence inversion, temporal supersession, span ungrounding, or monoculture defect.

7.2 Deterministic gate comparison (`G0` - `G5`)

Candidate Gate	TP	FP	TN	FN	False Acceptance Rate (FAR)	False Rejection Rate (FRR)	Overall Accuracy
`G0_gtp34_quorum_only` (`GTP-34` 4-family quorum only)	8	4,088	12,288	0	`24.96%` (`0.249634`)	`0.00%`	`75.05%` (`0.750488`)
`G1_cis35_local_single_doc` (Round 35 `CIS-35` on local primary doc)	2	126	16,250	6	`0.77%` (`0.007694`)	`75.00%` (`0.750000`)	`99.19%` (`0.991943`)
`G2_multidoc_attachment_no_prec_dag` (multi-doc closure, no `pi_prec` / `tau_exec`)	8	56	16,320	0	`0.34%` (`0.003420`)	`0.00%`	`99.66%` (`0.996582`)
`G3_multidoc_scalar_schema_rigid` (multi-doc + `pi_prec` + `tau_exec`, rigid scalar schema)	2	0	16,376	6	`0.00%` (`0.000000`)	`75.00%` (`0.750000`)	`99.96%` (`0.999634`)
`G4_ccs36_without_diversity` (`CCS-36` semantic gates without `≥ 3` families)	8	24	16,352	0	`0.15%` (`0.001466`)	`0.00%`	`99.85%` (`0.998535`)
`G5_ccs36_full` (`CCS-36` complete gate)	**8**	**0**	**16,376**	**0**	**`0.00%` (`0.000000`)**	**`0.00%` (`0.000000`)**	**`100.00%` (`1.000000`)**

7.3 Exact multi-hop reference-truncation failure surface (`E36-A`)

The table below reports the exact bundle-closure probability `P_closed(m, q) = q^(m - 1)`, induced common-mode truncation shock `rho_trunc(m, q) = 1 - q^(m - 1)`, and ungated 3-of-4 quorum failure probability `P_fail_quorum(m, q, p_indep = 0.15)` across document chain lengths `m in {1, 2, 3, 4, 5}` and per-hop retrieval recalls `q in {0.95, 0.85, 0.75}`:

Document Chain Length `m`	Per-Hop Recall `q`	Transitive Closure `q^(m-1)`	Induced Common-Mode Shock `rho_trunc`	Independent Quorum Tail (`p = 0.15`)	Ungated 3-of-4 Quorum Failure Probability
`m = 1` (0 hops)	`0.95`	`1.000000`	`0.000000`	`0.011981 (1.20%)`	`0.011981 (1.20%)`
`m = 2` (1 hop)	`0.95`	`0.950000`	`0.050000`	`0.011981 (1.20%)`	`0.061382 (6.14%)`
`m = 3` (2 hops)	`0.95`	`0.902500`	`0.097500`	`0.011981 (1.20%)`	`0.108313 (10.83%)`
`m = 4` (3 hops)	`0.95`	`0.857375`	`0.142625`	`0.011981 (1.20%)`	`0.152897 (15.29%)`
`m = 5` (4 hops)	`0.95`	`0.814506`	`0.185494`	`0.011981 (1.20%)`	`0.195253 (19.53%)`
`m = 1` (0 hops)	`0.85`	`1.000000`	`0.000000`	`0.011981 (1.20%)`	`0.011981 (1.20%)`
`m = 2` (1 hop)	`0.85`	`0.850000`	`0.150000`	`0.011981 (1.20%)`	`0.160184 (16.02%)`
`m = 3` (2 hops)	`0.85`	`0.722500`	`0.277500`	`0.011981 (1.20%)`	`0.286156 (28.62%)`
`m = 4` (3 hops)	`0.85`	`0.614125`	`0.385875`	`0.011981 (1.20%)`	`0.393233 (39.32%)`
`m = 5` (4 hops)	`0.85`	`0.522006`	`0.477994`	`0.011981 (1.20%)`	`0.484248 (48.42%)`
`m = 1` (0 hops)	`0.75`	`1.000000`	`0.000000`	`0.011981 (1.20%)`	`0.011981 (1.20%)`
`m = 2` (1 hop)	`0.75`	`0.750000`	`0.250000`	`0.011981 (1.20%)`	`0.258986 (25.90%)`
`m = 3` (2 hops)	`0.75`	`0.562500`	`0.437500`	`0.011981 (1.20%)`	`0.444239 (44.42%)`
`m = 4` (3 hops)	`0.75`	`0.421875`	`0.578125`	`0.011981 (1.20%)`	`0.583180 (58.32%)`
`m = 5` (4 hops)	`0.75`	`0.316406`	`0.683594`	`0.011981 (1.20%)`	`0.687385 (68.74%)`

8. Evidence Category 2: Live matched-compute multi-model benchmark (`E36-B`)

8.1 Witness architecture roster, reasoning configuration, and cryptographic execution ledger

`E36-B` executed `108` live witness evaluations across four distinct witness architecture families, recording every evaluation's prompt integrity receipt, latency, compositional section tuple, verbatim quote, and dangling-reference flag in the reproducibility archive.

Encevra Roster Identifier	Witness Architecture Family	Reasoning Configuration	Live Evaluations	Evaluation Arms Covered	Mean Latency (ms)	Receipt Verification
`witness-a` (Lead & Mono Quorum)	`Witness A` (Architecture Family A)	`reasoning: {"effort": "high"}` (Max reasoning)	54	`Arm 1` (12), `Arm 2` (48), `Arm 3/4` (12), `Arm 5` (12)	31,938.2	`54 / 54` SHA-256 verified
`witness-b` (Witness `W1`)	`Witness B` (Architecture Family B)	`reasoning: {"effort": "high"}`	18	`Arm 3/4` (6 partial + 6 complete), `Arm 5` (12)	15,333.7	`18 / 18` SHA-256 verified
`witness-c` (Witness `W2`)	`Witness C` (Architecture Family C)	`reasoning: {"effort": "high"}`	18	`Arm 3/4` (6 partial + 6 complete), `Arm 5` (12)	7,864.1	`18 / 18` SHA-256 verified
`witness-d` (Witness `W3`)	`Witness D` (Architecture Family D)	`reasoning: {"effort": "high"}`	18	`Arm 3/4` (6 partial + 6 complete), `Arm 5` (12)	9,001.2	`18 / 18` SHA-256 verified
Total (`E36-B`)	**4 Distinct Witness Architecture Families**	**High / Max Reasoning Enabled**	**108**	**All 5 Evaluation Arms (`M01-M12`)**	**21,335.6**	**`108 / 108` SHA-256 verified**

8.2 Frozen 12-packet multi-document enterprise benchmark (`M01 - M12`)

The `12` multi-document packets are partitioned into two strata of six packets each:

- **Stratum 1 (`S1_cross_document_dangling_override`, `M01-M06`):** In each packet, the primary instrument (`DOC-MSA`, `DOC-DPA`, `DOC-SLA`, or `DOC-LIC`) is 100% locally complete (`LOCALLY COMPLETE 100%`), so local single-document hash checking passes. However, the primary instrument incorporates a companion overriding instrument that is omitted from `B_partial` and attached in `B_complete`:
- `M01: DOC-MSA Section 8.1` (12-month / `1.5 million` general liability cap) superseded by `DOC-OF4 Section 3.2` (36-month / `4.5 million` Mission-Critical super-cap).
- `M02: DOC-DPA Section 5.1` (72-hour breach notification deadline) narrowed by `DOC-SBR Section 2.1` (12-hour Sovereign Banking Rider deadline; 30-hour notice breaches `DOC-SBR`).
- `M03: DOC-SLA Section 2.1` (`99.90%` standard monthly uptime commitment) elevated by `DOC-HAA Section 1.3` (`99.99%` High-Availability Addendum commitment; `99.94%` uptime triggers a 25% credit).
- `M04: DOC-MSA Section 11.1` (prohibition on subcontracting without standalone consent letter) reversed by `DOC-SCH-C Section 3.1` (pre-authorizing `AetherVault GmbH` in Frankfurt).

- `M05: DOC-MSA Section 9.2` (60-day non-renewal notice deadline) shortened by `DOC-CSL Section 1.1` (Closing Side Letter 15-day notice window; 20-day notice prevents automatic renewal).
- `M06: DOC-LIC Section 4.3` (100% royalty exemption for Disaster-Recovery Standby invocations) repealed and superseded by `DOC-AMD2 Section 2.1` (Second Amendment commercial reset making all Standby invocations billable at `0.04 units` per invocation).
- **Stratum 2 (`S2_precedence_dag_and_compositional_chains`, `M07-M12`):** Multiple instruments are attached simultaneously, testing Order-of-Precedence DAG resolution, specificity guardrails, compositional two-clause chains (directly testing Round 35's `C03` and `C12` hold bottlenecks), temporal amendment supersession, and unexecuted draft side-letter decoys:
- `M07`: Three-instrument precedence hierarchy (`DOC-DPA > DOC-OF2 > DOC-MSA` under `DOC-MSA Section 1.5`) where `DOC-DPA Section 9.3` (`5,000,000 unit` GDPR Article 82 cap) prevails over `DOC-OF2 Section 5.1` (`500,000 unit` general Order Form cap).
- `M08`: Precedence Specificity Guardrail (`DOC-MSA Section 1.3`) where `DOC-SOW3 Section 4.2` purports to assign all IP to Supplier but fails to cite `DOC-MSA Section 6.1` explicitly, leaving `DOC-MSA Section 6.1` controlling.
- `M09` (Round 35 `C03` Compositional Successor): Base Service Level Credit entitlement arises under `DOC-SLA Section 1.2` (`> 120` minutes Unscheduled Downtime) while the 45-minute quarterly HSM maintenance exclusion ceiling is defined in `DOC-EXB Section 3.4` (`540 - 45 = 495 > 120` minutes).
- `M10` (Round 35 `C12` Inline-Attestation Successor): Direct-competitor assignment prohibition under `DOC-MSA Section 15.2` (overriding `DOC-MSA Section 15.1`), paired with a cryptographically signed inline registry attestation (`INLINE-ATTESTED-EX-C`, `inline_fact_sufficiency = true`) stipulating that `Vanguard Kernels Inc.` is on Exhibit C.
- `M11`: Temporal amendment chronology across `DOC-MSA`, `DOC-AMD1` (First Amendment lifting the consequential damages waiver), and `DOC-AMD2` (active Second Amendment repealing `DOC-AMD1` and reinstating `DOC-MSA Section 8.2` waiver).
- `M12`: Executed multi-document binder (`DOC-MSA` + executed `DOC-OF7` placing exclusive venue in Zurich, Switzerland) containing an unsigned draft side letter decoy (`DOC-DRAFT-SL` proposing London, UK).

8.3 Aggregate comparison across the five evaluation arms

Evaluation Arm	Description	Correct Certified Releases	False Certified Releases	Fail-Closed Holds (`HOLD`)	False Certification Rate	Stratum 1 (`M01-M06`) False Releases
`Arm 1: Single-Witness-A-Max-Complete`	Single `witness-a` (`effort: high`) on complete multi-doc packet `B_complete`	`12 / 12` (`12 / 12` verified tuple & span)	`0 / 12`	`0 / 12`	`0.00%`	`0 / 6`
`Arm 2: CIS35-Local-Doc-3of4`	Round 35 `CIS-35` on locally complete primary doc (`B_partial` on `M01-M06`) + scalar schema	`4 / 12` (`33.33%`)	**`6 / 12`**	`2 / 12` (`M09, M10`)	**`50.00%`**	**`6 / 6 (100%)`**
`Arm 3: MultiFamily-Partial-Ungated`	`Witness A` + `Witness B` + `Witness C` + `Witness D` on `B_partial`, `GTP-34` only	`6 / 12` (`50.00%`)	**`6 / 12`**	`0 / 12`	**`50.00%`**	**`6 / 6 (100%)`**
`Arm 4: CCS-36 Pre-Reopen Gate`	4-family quorum gated by `CCS-36` on `B_partial` before cross-document bundle reopen	`6 / 12` (`50.00%`)	**`0 / 12`**	`6 / 12` (`M01-M06` held)	**`0.00%`**	**`0 / 6 (6/6 held)`**
`Arm 5: CCS-36 Complete Packet`	4-family quorum gated by `CCS-36` (`V_ref_closure * V_prec_dag * V_tuple * V_xspan`) on `B_complete`	**`12 / 12 (100.00%)`**	**`0 / 12 (0.00%)`**	**`0 / 12 (0.00%)`**	**`0.00%`**	**`0 / 6 (6/6 fixed)`**

8.4 Packet-by-packet live benchmark breakdown (`M01` - `M12`)

Packet	Stratum	Ground Truth (`Verdict`, `Base → Controlling Clause`, `Precedence`)	`Arm 1` (`Witness A Max, B_complete`)	`Arm 2` (`CIS-35` Local Single-Doc)	`Arm 3` (`MultiFamily` Partial, Ungated)	`Arm 4` (`CCS-36` on `B_partial`)	`Arm 5` (`CCS-36` `B_compl`
`M01`	`S1`	`FAIL` (`DOC-MSA 8.1 → DOC-OF4 3.2`, `DOC-OF4 > DOC-MSA`)	`FAIL` (tuple=True)	**`PASS (4-0 FALSE)`**	**`PASS (4-0 FALSE)`**	`HOLD_CCS36_DANGLING_REF_CLOSURE` (`4/4` flagged)	**`FAIL (4 tuples certifie`**
`M02`	`S1`	`FAIL` (`DOC-DPA 5.1 → DOC-SBR 2.1`, `DOC-SBR > DOC-DPA`)	`FAIL` (tuple=True)	**`PASS (4-0 FALSE)`**	**`PASS (4-0 FALSE)`**	`HOLD_CCS36_DANGLING_REF_CLOSURE` (`4/4` flagged)	**`FAIL (4 tuples certifie`**
`M03`	`S1`	`PASS` (`DOC-SLA 2.1 → DOC-HAA 1.3`, `DOC-HAA > DOC-SLA`)	`PASS` (tuple=True)	**`FAIL (4-0 FALSE)`**	**`FAIL (4-0 FALSE)`**	`HOLD_CCS36_DANGLING_REF_CLOSURE` (`4/4` flagged)	**`PASS (4 tuples certifie`**
`M04`	`S1`	`PASS` (`DOC-MSA 11.1 → DOC-SCH-C 3.1`, `DOC-SCH-C > DOC-MSA`)	`PASS` (tuple=True)	**`FAIL (4-0 FALSE)`**	**`FAIL (4-0 FALSE)`**	`HOLD_CCS36_DANGLING_REF_CLOSURE` (`4/4` flagged)	**`PASS (4 tuples certifie`**
`M05`	`S1`	`FAIL` (`DOC-MSA 9.2 → DOC-CSL 1.1`, `DOC-CSL > DOC-MSA`)	`FAIL` (tuple=True)	**`PASS (4-0 FALSE)`**	**`PASS (4-0 FALSE)`**	`HOLD_CCS36_DANGLING_REF_CLOSURE` (`4/4` flagged)	**`FAIL (4 tuples certifie`**

Packet	Stratum	Ground Truth (Verdict, Base → Controlling Clause, Precedence)	Arm 1 (Witness A Max, B_complete)	Arm 2 (CIS-35 Local Single-Doc)	Arm 3 (MultiFamily Partial, Ungated)	Arm 4 (CCS-36 on B_partial)	Arm 5 (CCS-36 B_compl
M06	S1	PASS (DOC-LIC 4.3 → DOC-AMD2 2.1, DOC-AMD2 > DOC-LIC)	PASS (tuple=True)	**FAIL (4-0 FALSE)**	**FAIL (4-0 FALSE)**	HOLD_CCS36_DANGLING_REF_CLOSURE (4/4 flagged)	**PASS (4 tuples certifie**
M07	S2	PASS (DOC-OF2 5.1 → DOC-DPA 9.3, DOC-DPA > DOC-OF2 > DOC-MSA)	PASS (tuple=True)	PASS (4-0 correct)	PASS (4-0 correct)	PASS (4/4 tuples, certified)	**PASS (4 tuples certifie**
M08	S2	PASS (DOC-SOW3 4.2 → DOC-MSA 6.1, DOC-MSA > DOC-SOW3)	PASS (tuple=True)	PASS (4-0 correct)	PASS (4-0 correct)	PASS (3/4 tuples, certified)	**PASS (3 tuples certifie**
M09	S2	PASS (DOC-SLA 1.2 → DOC-EXB 3.4, C03 Compositional Successor)	PASS (tuple=True)	**HOLD_CIS35_SCALAR_OR_EXHIBIT**	PASS (4-0 correct)	**PASS (4/4 tuples, certified)**	**PASS (4 tuples certifie**
M10	S2	PASS (DOC-MSA 15.1 → DOC-MSA 15.2, C12 Inline-Attestation Successor)	PASS (tuple=True)	**HOLD_CIS35_SCALAR_OR_EXHIBIT**	PASS (4-0 correct)	**PASS (4/4 tuples, certified)**	**PASS (4 tuples certifie**
M11	S2	PASS (DOC-MSA 8.2 → DOC-AMD2 1.1, DOC-AMD2 > DOC-AMD1 > DOC-MSA)	PASS (tuple=True)	PASS (4-0 correct)	PASS (4-0 correct)	PASS (4/4 tuples, certified)	**PASS (4 tuples certifie**
M12	S2	PASS (DOC-MSA 16.1 → DOC-OF7 9.1, DOC-DRAFT-SL decoy rejected)	PASS (tuple=True)	PASS (4-0 correct)	PASS (4-0 correct)	PASS (4/4 tuples, certified)	**PASS (4 tuples certifie**

9. Findings vs. interpretations (strict separation)

9.1 Measured findings (directly observed in E36-A and E36-B)

1. **Deterministic finding F36.1 (E36-A):** Across the 16,384 enumerated multi-document states (8 positive CCS_CERTIFIED states and 16,376 negative compromised states), G0_gtp34_quorum_only accepts 4,088 / 16,376 negative states (FAR = 24.96%), G1_cis35_local_single_doc accepts 126 / 16,376 negative states (FAR = 0.77%) while rejecting 6 / 8 positive states (FRR = 75.00%), G2_multidoc_attachment_no_prec_dag accepts 56 / 16,376 negative states (FAR = 0.34%), G3_multidoc_scalar_schema_rigid achieves FP = 0 but rejects 6 / 8 positive states (FRR = 75.00%), G4_ccs36_without_diversity accepts 24 / 16,376 negative states (FAR = 0.15%), and G5_ccs36_full achieves TP = 8, FP = 0, TN = 16,376, FN = 0 (FAR = 0.00%, FRR = 0.00%, Accuracy = 100.00%).
2. **Deterministic finding F36.2 (E36-A):** In the exact multi-hop reference-truncation surface with marginal independent witness error p_indep = 0.15 (1.20% independent 3-of-4 quorum failure), per-hop retrieval recall q = 0.85 elevates ungated 3-of-4 quorum failure to 16.02% at m = 2, 28.62% at m = 3, 39.32% at m = 4, and 48.42% at m = 5. At m = 3, per-hop recall must exceed q* = 0.964964 (96.50%) for an ungated 3-of-4 quorum to beat a single p_single = 0.08 full-packet checker.
3. **Empirical live-model finding F36.3 (E36-B):** On Stratum 1 (M01-M06), where each primary agreement (DOC-MSA, DOC-DPA, DOC-SLA, DOC-LIC) is 100% locally complete (V_scope_local = 1) but references an unattached companion instrument, all four homogeneous Witness A witnesses (Arm 2, 24 / 24 calls) and all four heterogeneous witness architecture families (Arm 3: Witness A, Witness B, Witness C, Witness D, 24 / 24 calls) emit unanimous 4-0 verdicts opposite to the complete multi-document ground truth (6 / 6 false certified releases on Stratum 1; 50.00% overall false certification rate across M01-M12).
4. **Empirical live-model finding F36.4 (E36-B):** Applying CCS-36 to the partial packets (Arm 4) reduces false certified releases from 6 / 12 (50.00%) to 0 / 12 (0.00%): on all 6 / 6 Stratum-1 packets (M01-M06), 4 / 4 witness families (24 / 24 calls) explicitly set unattached_external_reference = true and V_ref_closure blocks release fail-closed (HOLD_CCS36_DANGLING_REF_CLOSURE).
5. **Empirical live-model finding F36.5 (E36-B):** Once CCS-36 reopens the transitive multi-document bundle (Arm 5), all 6 / 6 Stratum-1 packets (M01-M06) flip from 4-0 false consensus to 4-0 verified correct consensus (24 / 24 calls), and all 6 / 6 Stratum-2 packets (M07-M12) achieve 4-0 correct verdict consensus with 4 / 4 (M07, M09, M10, M11, M12) and 3 / 4 (M08) verified compositional (base_entitlement_section, decisive_override_section) tuples. In particular, the compositional tuple and inline_fact_sufficiency predicate resolve both Round 35 hold archetypes (M09 and M10), raising certified correct releases from 10 / 12 (83.33% in Round 35 Arm 5 and 4 / 12 = 33.33% in Round 36 Arm 2) to 12 / 12 (100.00%) with 0 / 12 false releases and 0 / 12 false holds.

9.2 Interpretations and architectural implications

1. **Why single-document hash verification (CIS-35) is a special case of reference-graph closure (CCS-36):** Round 35 treated a contract as a single byte array D with digest H* = SHA256(D). In enterprise deployments, however, legal authority is distributed across a directed reference graph G_doc = (D, E_ref, pi_prec). A primary agreement that is 100% byte-complete locally (SHA256(D_1) verified) behaves identically to a truncated retrieval slice whenever an outgoing reference edge (D_1, D_k) in E_ref points to an unattached overriding instrument D_k. CCS-36 generalizes CIS-35 from single-document digest binding to transitive reference-graph closure (Cl(q, G_doc)).

2. **Why compositional clause-chain tuples eliminate false holds without weakening verification:** When a legal norm is structured as *"Rule A in Document 1 applies unless Condition B in Document 2 exceeds Ceiling C"*, asking four distinct reasoning architectures for a single scalar `governing_section` forces an arbitrary choice between naming the entitlement rule (`Rule A`) or the ceiling rule (`Ceiling C`). By factoring the receipt schema into `(base_entitlement_section, decisive_override_section, precedence_relation, inline_fact_sufficiency)`, `CCS-36` aligns the verification predicate with the actual two-node structure of legal override chains.

---

10. Scope limitations and threats to validity

1. **Sample size and benchmark boundary:** `E36-B` evaluated `12` multi-document enterprise packets (`M01-M12`) across `108` live witness evaluations (`5` evaluation arms). While the `6 / 6` (`100.00%`) false-consensus rate on unattached cross-document overrides (`M01-M06`) and the `12 / 12` (`100.00%`) recovery under `CCS-36` (`Arm 5`) are sharp, real enterprise data rooms can contain dozens of schedules and implicit cross-document definitions that do not name target instrument identifiers explicitly.
2. **Explicit vs. implicit incorporation edges:** In `M01-M06`, the primary instrument explicitly named the companion instrument (`DOC-OF4`, `DOC-SBR`, `DOC-HAA`, `DOC-SCH-C`, `DOC-CSL`, `DOC-AMD2`), enabling all four witness architectures to flag `unattached_external_reference = true` (`24 / 24` calls) on `B_partial`. When a later amendment modifies a base agreement *silently* (without the base agreement containing a forward pointer to that future amendment), inspecting the base agreement alone cannot reveal that a post-execution amendment exists unless the binder manifest or version ledger is queried.
3. **Research and production boundary:** Neither `E36-A` nor `E36-B` mutated any production Noetis service or customer workflow; all evaluations ran in an isolated research harness with deterministic SHA-256 receipt verification.

---

11. Next check (Round 37 handoff: `TDR-37`)

Round 36 Section 10.2 exposes the next structural frontier: **what happens when a base agreement `D_1` contains no forward pointer to a post-execution Amendment, Waiver, or Notice of Termination executed months later, or when asynchronous document stores serve stale binder snapshots across longitudinal contract lifecycles?** Round 37 should formulate and test **Temporal Drift and Post-Execution Amendment Supersession (`TDR-37`)**, pairing a bitemporal amendment-lineage diagnostic (`E37-A`) with a live multi-witness benchmark (`E37-B`) over longitudinal multi-amendment contract lifecycles where post-execution instruments silently repeal, pause, or reinstate earlier obligations without modifying the bytes of the base agreement.

---

12. References (`[S01]` - `[S12]`)

- **[S01]** John C. Knight and Nancy G. Leveson (1986). "An Experimental Evaluation of the Assumption of Independence in Multi-Version Programming." *IEEE Transactions on Software Engineering*, SE-12(1):96-109. https://doi.org/10.1109/TSE.1986.6312924
- **[S02]** Dave E. Eckhardt and Larry D. Lee (1985). "A Theoretical Basis for the Analysis of Multiversion Software Subject to Coincident Errors." *IEEE Transactions on Software Engineering*, SE-11(12):1511-1517. https://doi.org/10.1109/TSE.1985.231895
- **[S03]** Bev Littlewood and Douglas R. Miller (1995). "A Conceptual Model of Multi-Version Software." *Twenty-Fifth International Symposium on Fault-Tolerant Computing (FTCS-25 Highlights)*, pp. 150-159. https://doi.org/10.1109/FTCSH.1995.532633
- **[S04]** Dahlia Malkhi and Michael Reiter (1998). "Byzantine Quorum Systems." *Distributed Computing*, 11(4):203-213. https://doi.org/10.1007/s004460050050
- **[S05]** Flavio P. Junqueira and Keith Marzullo (2003). "Synchronous Consensus for Dependent Process Failures." *Proceedings of the 23rd International Conference on Distributed Computing Systems (ICDCS 2003)*, pp. 274-283. https://doi.org/10.1109/ICDCS.2003.1203476
- **[S06]** Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou (2023). "Self-Consistency Improves Chain of Thought Reasoning in Language Models." *International Conference on Learning Representations (ICLR 2023)*. https://arxiv.org/abs/2203.11171
- **[S07]** Yilun Du, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, and Igor Mordatch (2023). "Improving Factuality and Reasoning in Language Models through Multiagent Debate." *arXiv preprint arXiv:2305.14325*. https://arxiv.org/abs/2305.14325
- **[S08]** Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Shuming Shi, and Zhaopeng Tu (2023). "Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate." *arXiv preprint arXiv:2305.19118*. https://arxiv.org/abs/2305.19118
- **[S09]** Jie Huang, Xinyun Chen, Swaroop Mishra, Huaixiu Steven Zheng, Adams Wei Yu, Xinying Song, and Denny Zhou (2023). "Large Language Models Cannot Self-Correct Reasoning Yet." *arXiv preprint arXiv:2310.01798*. https://arxiv.org/abs/2310.01798
- **[S10]** Junyou Li, Qin Zhang, Yangbin Yu, Qiang Fu, and Deheng Ye (2024). "More Agents Is All You Need." *Transactions on Machine Learning Research (TMLR 2024)*. https://arxiv.org/abs/2402.05120
- **[S11]** Fan Wu, Emily Black, and Varun Chandrasekaran (2024). "Generative Monoculture in Large Language Models." *arXiv preprint arXiv:2407.02209*. https://arxiv.org/abs/2407.02209
- **[S12]** Mert Cemri, Melissa Z. Pan, Shuyi Yang, Lakshya A. Agrawal, Bhavya Chopra, Rishabh Tiwari, Kurt Keutzer, Aditya Parameswaran, Dan Klein, Kannan Ramchandran, Matei Zaharia, Joseph E. Gonzalez, and Ion Stoica (2025). "Why Do Multi-Agent LLM Systems Fail?" *arXiv preprint arXiv:2503.13657*. https://arxiv.org/abs/2503.13657