*Review*

# Secure and Privacy-Aware Networked Architectures for Monitoring Dual-Use Items and Technologies: A Scoping Review

**Eduard Mihăilescu *, Doru-Florin Chiper and Mihai Ghiuzan**

The supplementary package provides methodological, quantitative, corpus-level, claim-appraisal, and evidence-to-architecture information accompanying the review. Reference numbers correspond to the reference list in the main manuscript. A separate PRISMA-ScR checklist with reporting locations is submitted as Supplementary Materials.

## Corpus Selection and Composition

Table S1. Corpus selection and composition.

| Stage / corpus component | Records |
|---|---|
| Domain-mapping search | 70 source records |
| Mechanism-focused search | 30 distinct source records |
| Focused supplementary search (follow-up stage) | 20 source records |
| AI supply-chain search (October 2026) | 4 source records |
| Candidate source pool | 124 records |
| Records excluded before final inclusion | 34 records |
| Included substantive corpus | 90 sources |
| Peer-reviewed evidence | 28 sources |
| Technical standards and specifications | 22 sources |
| Legal and regulatory instruments | 10 sources |
| Institutional, operational, or vendor technical evidence | 30 sources |
| Methodological/reporting works (analytically separate) | 2 works |
| Final bibliography | 92 entries |

## Quantitative Implementation and Deployment Constraints

Table S2. Quantitative implementation evidence and deployment constraints.

| Mechanism / constraint | Platform or context | Reported metric / assessed failure | Evidence scope / proposed control |
|---|---|---|---|
| Hidden-DCT watermarking [9] | TSMC 90 nm ASIC and Xilinx Virtex-7; 512 × 512 image with a 128 × 128 watermark. | ASIC: 16,402 cycles, 7.053 μs at 2.32 GHz, 141,783 frames/s, and 103 GB/s. FPGA: 421.94 MHz; ASIC: 0.30498 mm² and 508.18 mW. FPGA: 6874 LUTs, 4922 flip-flops, and 754 mW. | PSNR 45.09 dB; NC ≥ 0.99 after JPEG QF 50. ASIC values are pre-layout synthesis results and FPGA power is estimated; findings are algorithm- and interface-specific. |
| FPGA watermark accelerator [53] | Pipelined FPGA implementation at 200 MHz. | Up to 3.2 GB/s; low-complexity add/shift realization; resource values remain device- and synthesis-specific. | Supports real-time processing. Direct area or power comparison with [9] is invalid because the algorithms, FPGA devices, and implementation conditions differ. |
| VRASED remote attestation [33] | OpenMSP430 on Basys3 Artix-7; 4 KiB of attested memory. | 3.6 million cycles; below 1 s at 8–25 MHz. Runtime grows linearly with the attested memory size; +122 LUTs (6.6%), +37 registers (5.4%), 6.3% total logic-cell increase, and approximately 4.5 KiB ROM. | Formal assurance and low hardware overhead are demonstrated for one low-end MCU architecture. Energy consumption and network latency were not measured. |
| C-FLAT control-flow attestation [34] | ARM TrustZone prototype; syringe-pump and PID-controller case studies. | PID-controller iteration overhead: 237 μs; approximately 0.03% when attested once per second. Total overhead grows linearly with control-flow events; maximum authenticator size in the PID case: 490 bytes; most runtime overhead arose in the Secure World algorithm. | Acceptability depends on the physical control-loop period and attestation frequency. The study does not provide a common area or power metric. |

| Mechanism / constraint | Platform or context | Reported metric / assessed failure | Evidence scope / proposed control |
|---|---|---|---|
| Merkle transparency proof [19] | RFC 9162 SHA-256 binary Merkle tree; authors' proof-size calculation. | Inclusion and consistency verification require O(log N) path processing; the hash payload is 640 bytes at $10^6$ entries and 960 bytes at $10^9$ entries, excluding framing, identifiers, signatures, and transport. | Proof size grows logarithmically; log storage, witness replication, and gossip traffic scale separately and do not establish input truth or event completeness. |
| Lifecycle-event storage model | Four events per day and 2 KiB per event, before indexes, raw evidence, WORM copies, or replication. | 2.85 MiB per asset-year; 27.85 GiB/year for 10,000 assets; 278.47 GiB/year for 100,000 assets; 2.72 TiB/year for 1,000,000 assets. | Illustrative capacity calculation rather than an empirical benchmark. Retention, replication, indexing, and evidence bundles can multiply the result. |
| Logic area | PUF interfaces, secure key storage, attestation roots, watermarking datapaths, cryptographic accelerators | Protection logic displaces application resources, prevents timing closure, or enlarges the trusted hardware boundary | Use the smallest, defensible root of trust; reuse datapaths only where isolation remains demonstrable; report technology, synthesis stage, clock, and occupied resources; reserve margin for routing and integration |
| Power and energy | Continuous watermarking, periodic or collective attestation, protected memory access, secure communication | Battery depletion, thermal throttling, reduced attestation frequency, or disabled monitoring during peak workload | Define energy per evidence operation and duty cycle; use event-triggered or adaptive frequency; batch non-urgent work; retain a protected low-power checkpoint path |
| Memory and storage | Attestation code, helper data and error correction, event logs, reference material, evidence buffers, signatures and receipts | Incomplete event capture, truncated evidence, inability to reconstruct state, or exposure caused by unprotected spill storage | Bound local buffers; protect helper data and cryptographic state; checkpoint compact digests; move raw evidence to authenticated external storage; preserve overflow and loss indicators |

| Mechanism / constraint | Platform or context | Reported metric / assessed failure | Evidence scope / proposed control |
|---|---|---|---|
| Latency | Quote generation, secure-world transitions, control-flow accumulation, event-log reconstruction, policy and receipt processing | Evidence becomes stale, an authorization decision arrives after the relevant operation, or a real-time deadline is missed | Separate local admission from deferred institutional verification; pre-position reference material; bind short-lived results to the session; report end-to-end and component latency |
| Bandwidth and communication cost | Evidence bundles, event logs, endorsements, revocation data, collective-attestation messages, transparency receipts and witness traffic | Congestion, denial of service, missing reports, or inability to operate through intermittent links | Transmit digests and deltas where sufficient; batch and hierarchically aggregate; use bounded offline queues; record communication gaps and reconcile them after reconnection |
| Throughput | Streaming watermark insertion or detection, high-rate sensor evidence, event ingestion and registry anchoring | Markers or events are dropped, sampling is reduced, or the evidence path becomes the application bottleneck | Use pipelined or parallel datapaths when line-rate operation is required; dimension ingestion and checkpoint frequency from peak rather than average load; apply backpressure without silently discarding evidence |
| Determinism and jitter | Hard real-time control, control-flow attestation, shared accelerators, caches, interrupts, network and remote-verifier dependencies | Variable worst-case execution time violates a safety or control-loop bound even when average latency is acceptable | Perform worst-case execution-time and jitter analysis; use bounded evidence profiles and paths; isolate nondeterministic services from hard real-time loops; define the permitted degraded mode when the bound cannot be met |

| Mechanism / constraint | Platform or context | Reported metric / assessed failure | Evidence scope / proposed control |
|---|---|---|---|
| Resilience | Local roots of trust, counters, logs, offline operation, verifier and registry services, witnesses and federation links | Reset, rollback, power loss, partition, partial quorum failure, or component replacement creates an unresolvable evidentiary gap | Use monotonic counters and boot epochs, anti-rollback state, atomic protected checkpoints, redundant verification and witness paths, explicit fail-safe or fail-operational policy, and accountable post-partition reconciliation |
| Atlas ML provenance[24] | BERT fine-tuning; Intel TDX prototype on Intel Xeon Gold 5520+; CPU-only PyTorch pipeline. | Training overhead below 8%; 8 KB C2PA manifests; up to 50% verification latency reduction with caching. | Preliminary platform-specific results; provenance integrity and receipt verification do not demonstrate dataset truth or end-use authorization. |

Basis: Published measurements are cited; proof-size and storage calculations, risk scenarios, and proposed controls are authors' work. Abbreviations in Table S2: DCT, discrete cosine transform; ASIC, application-specific integrated circuit; FPGA, field-programmable gate array; LUT, lookup table; PSNR, peak signal-to-noise ratio; NC, normalized correlation; JPEG, Joint Photographic Experts Group; QF, quality factor; PID, proportional–integral–derivative; WORM, write once, read many; MCU, microcontroller unit; ROM, read-only memory.

## Source-Level Corpus Characteristics

Table S3 summarizes the classification and lifecycle mapping of all 90 substantive sources.

The October 2026 AI-focused proceedings search identified four eligible peer-reviewed AI supply-chain studies[23–26] addressing executable-model attack pathways, attested ML lifecycle evidence, and model-release verification. The August PubMed and IEEE Xplore retrieval counts are reported separately.

Table S3. Source-level classification and analytical mapping of the 90-source technical corpus.

| Ref. | Source type | Primary focus | Authors' lifecycle mapping |
|---|---|---|---|
| [12] | Legal/regulatory | EU dual-use authorization and transfer control | L6 Policy/verification |
| [13] | Legal/regulatory | US dual-use classification, licensing, and controls | Cross-layer / contextual |
| [75] | Institutional/operational/vendor | Cybersecurity supply-chain risk management | Cross-layer / contextual |
| [76] | Standard/specification | Counterfeit and maliciously tainted product assurance | Cross-layer / contextual |

| Ref. | Source type | Primary focus | Authors' lifecycle mapping |
|---|---|---|---|
| [1] | Peer-reviewed study | PUF foundations and hardware identity | L2 Identity/trust anchors |
| [78] | Peer-reviewed study | PUF enrollment, reliability, and key derivation | L2 Identity/trust anchors |
| [14] | Peer-reviewed study | PUF integration in IoT systems | L2 Identity/trust anchors |
| [39] | Peer-reviewed study | Watermarking robustness and trade-offs | L4 Provenance/lifecycle events |
| [54] | Peer-reviewed study | DCT-domain persistent tagging | L1/L4 Observation/provenance |
| [55] | Peer-reviewed study | DCT-domain persistent tagging | L1/L4 Observation/provenance |
| [56] | Peer-reviewed study | Image watermark robustness | L4 Provenance/lifecycle events |
| [9] | Peer-reviewed study | Hardware watermark implementation | L4 Provenance/lifecycle events; L1/L4 Observation/provenance |
| [53] | Peer-reviewed study | High-throughput hardware watermarking | L4 Provenance/lifecycle events; L1/L4 Observation/provenance |
| [57] | Peer-reviewed study | Transform selection for watermarking | L4 Provenance/lifecycle events |
| [10] | Peer-reviewed study | Transform hardware and obfuscation | Cross-layer / contextual |
| [77] | Peer-reviewed study | Transform implementation efficiency | L1/L4 Observation/provenance |
| [11] | Institutional/operational/vendor | Strategic-trade control context | Cross-layer / contextual |
| [40] | Legal/regulatory | Cyber-surveillance export authorization | L6 Policy/verification |
| [47] | Institutional/operational/vendor | Diversion prevention and end-use verification | L6 Policy/verification; L7 Gov./audit/feedback |
| [46] | Legal/regulatory | Firearms marking, tracing, and trafficking control | L7 Gov./audit/feedback |
| [5] | Institutional/operational/vendor | Nuclear safeguards verification | L6 Policy/verification; L7 Gov./audit/feedback |
| [48] | Institutional/operational/vendor | Software component transparency and provenance | L4 Provenance/lifecycle events; L5 Registry/transparency |
| [38] | Standard/specification | Signed digital-content provenance | L4 Provenance/lifecycle events |
| [51] | Institutional/operational/vendor | Manufacturing supply-chain traceability | L4 Provenance/lifecycle events |
| [41] | Institutional/operational/vendor | Field reconstruction of component diversion | L7 Gov./audit/feedback |
| [18] | Institutional/operational/vendor | Field reconstruction of UAV component diversion | L7 Gov./audit/feedback |
| [29] | Legal/regulatory | Item-level identity and lifecycle records | L4 Provenance/lifecycle events |
| [37] | Standard/specification | Chain-of-custody terminology and models | Cross-layer / contextual |
| [2] | Standard/specification | Event-based supply-chain traceability | L4 Provenance/lifecycle events |
| [69] | Institutional/operational/vendor | Chemical declarations and reconciliation | L7 Gov./audit/feedback |
| [58] | Legal/regulatory | Multilateral dual-use and munitions classifications | Cross-layer / contextual |

| Ref. | Source type | Primary focus | Authors' lifecycle mapping |
|---|---|---|---|
| [59] | Legal/regulatory | Nuclear-related dual-use supplier controls | Cross-layer / contextual |
| [62] | Institutional/operational/vendor | Life-sciences dual-use governance | L7 Gov./audit/feedback |
| [63] | Legal/regulatory | Dual-use research and pathogen oversight | Cross-layer / contextual |
| [89] | Institutional/operational/vendor | Ammunition lifecycle, inventory, and tracing | L4 Provenance/lifecycle events |
| [87] | Institutional/operational/vendor | Commercial cyber-intrusion governance | L7 Gov./audit/feedback |
| [4] | Legal/regulatory | US UAS remote-identification requirements | Cross-layer / contextual |
| [85] | Standard/specification | Remote ID and tracking protocol requirements | L2 Identity/trust anchors |
| [71] | Institutional/operational/vendor | Dual-use foundation-model misuse risk | Cross-layer / contextual |
| [70] | Institutional/operational/vendor | Secure software development for Generative AI | Cross-layer / contextual |
| [3] | Standard/specification | Unique identification of individual products | Cross-layer / contextual |
| [64] | Institutional/operational/vendor | Internal compliance and transaction screening | Cross-layer / contextual |
| [43] | Institutional/operational/vendor | Permissioned declaration reconciliation | L7 Gov./audit/feedback |
| [52] | Legal/regulatory | Lifecycle cybersecurity for products with digital elements | L4 Provenance/lifecycle events |
| [90] | Institutional/operational/vendor | Frontier-model capability and risk commitments | Cross-layer / contextual |
| [72] | Institutional/operational/vendor | Space surveillance, cataloguing, and tracking | L7 Gov./audit/feedback |
| [66] | Institutional/operational/vendor | Space-domain evidence and data requirements | L7 Gov./audit/feedback |
| [45] | Institutional/operational/vendor | Supply-chain traceability meta-framework | L4 Provenance/lifecycle events |
| [92] | Peer-reviewed study | Distributed governance of dual-use research | L7 Gov./audit/feedback |
| [60] | Institutional/operational/vendor | Long-term space sustainability | L7 Gov./audit/feedback |
| [61] | Institutional/operational/vendor | Space situational awareness and evidence sharing | L7 Gov./audit/feedback |
| [91] | Institutional/operational/vendor | Non-state acquisition and weaponization of UAS | Cross-layer / contextual |
| [65] | Peer-reviewed study | Cyber-surveillance regulation | L6 Policy/verification |
| [88] | Peer-reviewed study | Spyware and dual-use export control | L6 Policy/verification |
| [44] | Institutional/operational/vendor | European UAS remote-identification implementation | Cross-layer / contextual |
| [67] | Standard/specification | European direct remote-identification technical standard | Cross-layer / contextual |
| [6] | Standard/specification | Remote-attestation roles and trust model | L3 Attestation/evidence |

| Ref. | Source type | Primary focus | Authors' lifecycle mapping |
|---|---|---|---|
| [7] | Standard/specification | Attestation claims and token encoding | L3 Attestation/evidence |
| [30] | Standard/specification | TPM attestation, measured boot, and key protection | L3 Attestation/evidence |
| [80] | Institutional/operational/vendor | Intel SGX DCAP attestation and collateral | L3 Attestation/evidence |
| [81] | Institutional/operational/vendor | AMD confidential-VM attestation | L3 Attestation/evidence |
| [82] | Institutional/operational/vendor | Arm Realm and platform attestation | L3 Attestation/evidence |
| [84] | Standard/specification | Operator and authenticator public-key credentials | L2 Identity/trust anchors |
| [21] | Standard/specification | Semantic provenance model | L4 Provenance/lifecycle events |
| [22] | Peer-reviewed study | Authorized software supply-chain steps | Cross-layer / contextual |
| [50] | Standard/specification | Build provenance and supply-chain integrity | L4 Provenance/lifecycle events |
| [73] | Standard/specification | System and software package data exchange | Cross-layer / contextual |
| [68] | Standard/specification | Component, service, and cryptographic BOMs | Cross-layer / contextual |
| [49] | Peer-reviewed study | Keyless signing and transparency | L5 Registry/transparency |
| [32] | Standard/specification | Transparency services and verifiable receipts | L5 Registry/transparency |
| [19] | Standard/specification | Merkle inclusion and consistency proofs | L5 Registry/transparency |
| [42] | Standard/specification | Signed, sequenced, and replay-resistant logging | L5 Registry/transparency |
| [35] | Standard/specification | Issuer–holder–verifier credential model | L2 Identity/trust anchors |
| [8] | Standard/specification | Selective disclosure for JSON Web Tokens | L6 Policy/verification |
| [74] | Standard/specification | Selective and unlinkable derived proofs | L6 Policy/verification |
| [20] | Peer-reviewed study | Privacy-preserving hardware attestation | L3 Attestation/evidence; L6 Policy/verification |
| [33] | Peer-reviewed study | Formally verified embedded remote attestation | L3 Attestation/evidence; L1/L4 Observation/provenance |
| [34] | Peer-reviewed study | Control-flow attestation | L3 Attestation/evidence |
| [17] | Peer-reviewed study | Forward-integrity audit logging | L5 Registry/transparency |
| [86] | Peer-reviewed study | Collective and swarm attestation | L3 Attestation/evidence |
| [36] | Standard/specification | Workload identity and trust-domain federation | Cross-layer / contextual |
| [16] | Peer-reviewed study | SGX transient-execution attack | L3 Attestation/evidence |
| [83] | Peer-reviewed study | SEV-ES/SNP fault-injection attack | L3 Attestation/evidence |
| [31] | Institutional/operational/vendor | Attribute-based authorization and information sharing | L6 Policy/verification |
| [79] | Institutional/operational/vendor | Cryptographic access control using attributes | L6 Policy/verification |
| [15] | Institutional/operational/vendor | ZKP, MPC, and privacy/public-auditability mechanisms | L6 Policy/verification |

| Ref. | Source type | Primary focus | Authors' lifecycle mapping |
|---|---|---|---|
| [23] | Peer-reviewed study | MalHug model-hub code poisoning, malicious models and loading scripts | L1/L4 Artifact observation and provenance; cross-layer security |
| [24] | Peer-reviewed study | Atlas ML checkpoint lineage, TEE attestations and transparency records | L3 Attestation/evidence; L4 provenance; L5 registry |
| [25] | Peer-reviewed study | Laminator attestable ML property cards for training and inference | L3 Attestation/evidence; L6 verification |
| [26] | Peer-reviewed study | ML publisher signing and dataset-property supply-chain claims | L2 identity; L4 provenance; L5 transparency |

Source: Authors' classification of the included sources. Abbreviations in Table S3: UAS, unmanned aircraft system; TPM, Trusted Platform Module; SGX, Software Guard Extensions; SEV-ES, Secure Encrypted Virtualization–Encrypted State; SNP, Secure Nested Paging; VM, virtual machine; ZKP, zero-knowledge proof; MPC, multiparty computation.

## Search Strategy, Eligibility, and Source Assessment

Table S4. Data-charting and coding fields.

| Coding field | Operational categories / values |
|---|---|
| Source metadata | Citation key; short label; source type; repository or route; version; publication status; maintenance authority. |
| Monitored entity | Physical item; device; material; software; firmware; AI model; model weights/checkpoint; dataset; content; data stream; cyber-physical subsystem. |
| Identity and binding | Serial/UII; credential; certificate; PUF; secure element; TPM/TEE identity; physical-to-digital binding; ownership or lifecycle transition. |
| Attestation | Attester; Verifier; Relying Party; Evidence format; nonce/freshness; key binding; endorsements; reference values; TCB/collateral; appraisal result. |
| Provenance and lifecycle state | Watermark; hash; signature; W3C PROV; SBOM; in-toto; SLSA; C2PA; ML-BOM; fine-tuning checkpoint; transformation, transfer, integration, maintenance, and incident events. |
| Registry and transparency | Central; federated; permissioned; event hub; EPCIS; SCITT; Merkle log; receipt; checkpoint; witness; gossip/cross-logging; correction provenance. |
| Security and privacy | Authentication; authorization; ABAC; attribute-based encryption; purpose; role; environment; minimization; pseudonymization; selective disclosure; ZKP; MPC; confidential computing; linkability; output leakage; retention; deletion; accountable access. |
| Verification and policy | Integrity checks; reference-value appraisal; policy identifier and hash; decision class; reason code; escalation; reproducibility; human review. |
| Interoperability | Schema; media type; profile; trust anchor; trust bundle; credential semantics; version negotiation; reconciliation procedure. |
| Lifecycle governance | Revocation; key rotation; policy update; incident feedback; audit rights; dispute handling; institutional responsibility. |

| Coding field | Operational categories / values |
|---|---|
| Constraints and failure modes | Area; energy; memory; latency; throughput; bandwidth; determinism; replay; stale collateral; detachment; omission; equivocation; schema mismatch; excessive disclosure. |
| Interpretive scope | Normative; empirical; operational; legal; vendor-specific; domain-specific; provisional specification; residual uncertainty. |

Source: Authors' coding framework. Abbreviations in Table S4: UII, unique item identifier; TEE, trusted execution environment; TCB, trusted computing base; SBOM, software bill of materials; ABAC, attribute-based access control; PROV, W3C provenance framework; C2PA, Coalition for Content Provenance and Authenticity; SLSA, Supply-chain Levels for Software Artifacts.

During source assessment, E.M., D.-F.C. and M.G. undertook cross-author verification at selection, charting and analytical coding. In this collaborative procedure, a co-author checked a recorded judgment independently against its primary source and the shared criteria before the entry was consolidated by the team. At selection, eligibility, relevance, source authority and version-precedence decisions were cross-checked against full texts and the predefined operational rules; questions requiring adjudication were resolved by revisiting the supporting passages. The final inclusion decisions were reconciled with the 90-source inventory in Table S3. At charting, another author checked the recorded attributes against the source document under the fields in Table S4, including evidence objects, trust assumptions, lifecycle events, verification context, deployment constraints and claim scope. At coding, co-authors checked one another's interpretations of assurance functions, evidence strength, lifecycle-layer assignment and boundary failure modes against cited evidence and common definitions. Divergent readings were discussed and reconciled before the synthesis was finalized. Reference-identifier reconciliation verified the source references used in the claim appraisals and architecture mappings against the included inventory.

Table S5. Search query families and source routes.

| Search family / route | Search logic or repository route |
|---|---|
| Cross-architecture core | ("dual-use" OR "dual use") AND (monitor* OR traceab* OR provenance OR attestation OR "root of trust" OR registry OR audit*) AND (architectur* OR network* OR embedded OR "cyber-physical") |
| Attestation and hardware trust | ("remote attestation" OR RATS OR EAT OR TPM OR TEE OR "confidential computing" OR PUF) AND (Evidence OR Verifier OR "reference values" OR collateral OR revocation OR freshness OR "key binding") |
| Provenance and supply chain | (provenance OR SBOM OR "software bill of materials" OR in-toto OR SLSA OR C2PA OR Sigstore OR watermark*) AND (lifecycle OR traceab* OR "supply chain" OR event OR registry) |

| Search family / route | Search logic or repository route |
|---|---|
| Privacy-preserving verification | ("selective disclosure" OR "verifiable credentials" OR SD-JWT OR BBS OR "direct anonymous attestation" OR "attribute-based access control" OR ABAC OR "attribute-based encryption" OR ABE OR "zero-knowledge" OR ZKP OR "secure multiparty computation" OR MPC OR "confidential computing" OR pseudonym*) AND (verification OR attestation OR provenance OR audit OR linkability OR authorization OR correlation) |
| Registry, secure logging, and transparency | ("secure logging" OR "forward integrity" OR SCITT OR transparency OR "Merkle log" OR append-only OR witness* OR gossip OR equivocation) AND (audit OR provenance OR registry OR evidence) |
| Embedded/VLSI constraints | (attestation OR watermark* OR PUF) AND (FPGA OR VLSI OR IoT OR embedded OR "resource-constrained") AND (latency OR area OR energy OR throughput OR memory OR "communication cost") |
| Normative repositories | IETF RATS and SCITT; W3C PROV, VC, WebAuthn, and BBS; TCG TPM; ISO/IEC; GS1; ASTM; CEN; C2PA; SPDX; CycloneDX; SLSA; SPIFFE |
| Legal and institutional repositories | EUR-Lex; U.S. EAR/eCFR/BIS; Wassenaar; NSG; NIST; IAEA; OPCW; SIPRI; UNIDIR; CAR; ESA; IISS; UNOCT; FAA/EASA; Intel, AMD, and Arm technical portals |
| Citation tracing | Backward and forward citation tracing from foundational PUF, watermarking, attestation, secure-logging, software-supply-chain, and privacy sources; authoritative versions selected where multiple editions existed. |
| Peer-reviewed publisher / proceedings routes | In addition to the documented IEEE Xplore database search, relevant papers were located by manual examination of ACM, USENIX, Springer Nature, Elsevier, and MDPI publication or proceedings pages and by backward/forward citation tracing. Reproducible database query strings are specified separately in Table S6. |
| Search-wave structure | Two principal waves were undertaken: (1) broad domain mapping and (2) mechanism-focused lifecycle-control searches using the thematic families above. A focused supplementary search identified 20 candidate records (Table S1). August screening and charting took place during 1–31 August 2026, including research sessions on 5, 12, 19 and 26 August 2026. |

| Search family / route | Search logic or repository route |
|---|---|
| Supplementary route assessment | Manually located papers, authoritative texts, and citation-chain discoveries were assessed using the same eligibility criteria in Table S7 and entered in the classified source inventory in Table S3. |
| AI supply-chain proceedings search | Publisher proceedings and full-paper search for ("machine learning model" OR "pre-trained model") AND ("provenance" OR "malicious code" OR "attestation" OR "model signing"); four relevant papers retained[23–26]. These records were assessed separately from the August PubMed and IEEE Xplore searches. |

Source: Authors' search strategy. Table S5 records thematic search families and manual publisher/proceedings discovery. Table S6 gives the exact PubMed and IEEE Xplore database query syntax.

Abbreviations in Table S5 (technical and standards routes): IETF, Internet Engineering Task Force; VC, verifiable credential; BBS, Boneh–Boyen–Shacham; TCG, Trusted Computing Group; VLSI, very-large-scale integration; IoT, Internet of Things; SD-JWT, selective disclosure for JSON Web Tokens; ABE, attribute-based encryption; SPDX, Software Package Data Exchange; SPIFFE, Secure Production Identity Framework for Everyone; CEN, European Committee for Standardization.

Institutional-route abbreviations: EAR, Export Administration Regulations; eCFR, electronic Code of Federal Regulations; BIS, Bureau of Industry and Security; NSG, Nuclear Suppliers Group; NIST, National Institute of Standards and Technology; OPCW, Organisation for the Prohibition of Chemical Weapons; SIPRI, Stockholm International Peace Research Institute; UNIDIR, United Nations Institute for Disarmament Research; CAR, Conflict Armament Research; ESA, European Space Agency; IISS, International Institute for Strategic Studies; UNOCT, United Nations Office of Counter-Terrorism; FAA, Federal Aviation Administration; EASA, European Union Aviation Safety Agency.

Table S6. Database-specific reproducibility strategies.

| Database / search wave | Exact reproducibility strategy | Limits and coverage |
|---|---|---|
| PubMed Domain-mapping wave | Field tags: [tiab] = Title/Abstract; [lang] = Language. (("dual-use"[tiab] OR "dual use"[tiab]) AND (monitor*[tiab] OR track*[tiab] OR traceab*[tiab] OR "end-use"[tiab] OR "end user"[tiab] OR "export control"[tiab] OR governance[tiab] OR "cyber-surveillance"[tiab] OR spyware[tiab])) AND english[lang] | English; no lower year limit; coverage through August 2026. |

| Database / search wave | Exact reproducibility strategy | Limits and coverage |
|---|---|---|
| IEEE Xplore Attestation / hardware-trust wave | IEEE Xplore Advanced Search; both boxes = All Metadata. Box 1: "remote attestation" OR RATS OR EAT OR TPM OR TEE OR "confidential computing" OR PUF AND Box 2: Evidence OR Verifier OR "reference values" OR collateral OR revocation OR freshness OR "key binding" | All publication years; coverage through August 2026. |

The PubMed and IEEE Xplore searches were conducted during 1–31 August 2026, including research sessions on 5, 12, 19 and 26 August 2026. The corresponding query syntax appears in Table S6, and the AI-focused proceedings search is documented in Tables S1 and S5.

Source: Authors' search strategy. Table S6 provides executable search strings for the two recorded August database searches. Publisher and proceedings routes listed in Table S5, together with backward and forward citation tracing, supported source identification and follow-up. Platform-specific Boolean syntax and retrieval counts are available only for PubMed and IEEE Xplore.

Table S7. Operational inclusion and exclusion criteria.

| Criterion | Operational rule |
|---|---|
| Inclusion 1: identifiable mechanism | Source describes a mechanism, interface, role, evidence format, workflow, attack, constraint, or verification process. |
| Inclusion 2: architectural relevance | Source maps to at least one lifecycle layer and one assurance, privacy, interoperability, or deployment dimension. |
| Inclusion 3: evidentiary proximity | Source type is appropriate to the claim: technical study, standard, legal instrument, operational record, or vendor evidence structure. |
| Inclusion 4: stable and verifiable text | Stable full text, DOI, authoritative specification, official repository, or reproducible institutional citation is available. |
| Exclusion 1: generic commentary | High-level cybersecurity, AI ethics, geopolitical, or policy commentary lacks a monitorable mechanism or verification function. |
| Exclusion 2: non-authoritative duplicate | A later or more authoritative version preserves the relevant mechanism. |
| Exclusion 3: commercial description | Product material lacks reproducible evidence structure, protocol detail, or independent corroboration. |
| Exclusion 4: insufficient chartability | Document is too limited to code for architectural function, evidence object, trust assumption, or failure mode. |
| Conditional inclusion: provisional specification | Candidate Recommendation, Internet-Draft, or work item is directly relevant to maturity or research-gap analysis. |
| Eligibility scope: language | English-language source or authoritative English-language version available for source-level charting. |
| Eligibility scope: temporal coverage | No lower year cutoff; sources available through October 2026. |

| Criterion | Operational rule |
|---|---|
| Eligibility scope: publication/status | Peer-reviewed publications, authoritative standards/specifications, legal/regulatory instruments, and inspectable institutional/operational/vendor technical sources were eligible. Provisional standards were eligible only when directly relevant to maturity or research-gap analysis and were identified as provisional. |

Source: Authors' eligibility criteria.

## Claim-Level Evidence Assessment and Reproducibility

The Strong, Moderate, and Limited ratings apply to explicitly bounded propositions and are separate from the source-type categories in Table S3. Strong is appropriate for directly specified role, interface, format or data-model semantics within an authoritative document's stated scope, for binding legal obligations within their remit, or for convergent comparable empirical results from independent primary evidential lines. Moderate refers to relevant empirical or operational findings whose application remains context-specific. Limited indicates indirect, provisional, or system-level extrapolation without corresponding direct validation. A claim about a standard's prescribed data model and a claim about the completeness of real-world records require separate appraisal. A reference-identifier reconciliation of the source mappings confirmed 90 unique substantive entries in Table S3, distinct from methodology references [27,28]; every citation used in the evidence-assessment and architecture-mapping tables maps to an included substantive reference. Sources with a shared primary dataset, measurement, or underlying normative assertion were counted as a single evidential line for the claim concerned.

Table S8. Operational evidence-strength criteria and representative claim-level applications.

| Rating / decision rule | Illustrative claim and reference | Inference boundary |
|---|---|---|
| Strong: directly defined roles, interfaces or formats in an authoritative document, or binding legal obligations, within scope | RATS specifies attestation roles [6], and EAT specifies structured attestation claims [7]. | The rating applies to defined interface and role semantics, not to field effectiveness. |
| Strong: directly specified transparency and record-integrity semantics | SCITT defines signed-statement registration and receipt semantics [32]; Certificate Transparency specifies inclusion and consistency proofs [19]. | These documents establish defined registration or consistency properties; submitted input may still be incomplete or false. |
| Strong: normative lifecycle and custody data-model semantics | EPCIS defines lifecycle-event representations [2]; ISO 22095 specifies chain-of-custody terminology and models [37]. | Strong applies to the defined model and its semantics; real-world event coverage and custody accuracy require separate operational evidence. |

| Rating / decision rule | Illustrative claim and reference | Inference boundary |
|---|---|---|
| Moderate: primary empirical demonstration bounded by the studied platform and threat model | VRASED reports verified remote-attestation properties and prototype constraints [33]; C-FLAT evaluates control-flow attestation in defined scenarios [34]. | Claims are limited to the stated hardware, instrumented workflow, measurements and attacker assumptions. |
| Moderate: documented operational case or incident reconstruction with limited transferability | Conflict Armament Research documents component traces in specific field investigations [18,41]. | The documented cases support reconstruction within their evidence scope; no general incident rate or monitoring effectiveness is inferred. |
| Limited: cross-mechanism failure hypothesis supported indirectly by distinct mechanism families | Possible loss of continuity among identity binding, lifecycle-event capture and registered statements [1,2,32]. | This is the authors' comparative boundary inference; its frequency and detection effectiveness have not been measured as an integrated system. |
| Limited: original end-to-end architectural synthesis and proposed mitigation | The seven-layer decomposition and interface controls proposed in Section 3.4. | The source corpus supports the component functions; integrated operation and performance of the complete architecture remain to be validated. |
| Moderate: documented AI-platform studies with bounded implementation results | MalHug screened 705K+ models/176K+ datasets[23]; Atlas measured overhead in one ML pipeline [24]; Laminator evaluated attested ML property cards [25]. | Findings apply to their stated platform, sampling, TEE and attack assumptions; they do not validate a shared seven-layer deployment. |
| Limited: proposed cross-system AI model-release hand-off | Publisher signatures and dataset-property claims discussed in the ICML position analysis[26], with CycloneDX model/dataset metadata [68]. | Treating signatures, model-version provenance, receipt and recipient policy as one complete release control is the authors' synthesis. |

Source: Authors' claim-level and architectural synthesis of the sources cited in the main manuscript.

## Evidence-to-Architecture Mapping

The table relates each proposed logical layer to representative sources, the directly supported evidentiary function, and the authors' separation rationale. Source identifiers correspond to the manuscript reference list; Table S3 continues to provide the complete source-level classification of the 90-source corpus. A single source may support several functions. The mapping supports a logical architecture rather than a requirement for seven independently deployed components.

Studies of AI model security, provenance, attestation, and signing provide cross-domain evidence for checkpoint identity, execution-attested properties, model lineage, publisher signatures and registry verification[23–26]. Their roles are incorporated in the representative-source mapping;

where a downstream hand-off check is not implemented by a cited system, it remains an author-proposed requirement.

Table S9. Evidence-to-architecture mapping and justification for the seven logical layers.

| Layer; representative references | Source-supported function | Authors' layer boundary and design decision |
|---|---|---|
| L1. Monitored entity and observation. [3,4,29,44,85] | Identifiable items and operational observation schemes; unique item identification and Remote ID concern the observed entity. | Define the observation and entity being asserted before validating credentials or history. The entity-to-observation link remains separately checkable. |
| L2. Identity and trust anchors. [1,14,26,35,36,78,84] | PUF-based identity and enrollment constraints; credentials, authenticators and federated workload identities. | Keep entity and credential binding separate from point-in-time state appraisal; key compromise and ownership change affect this trust relation. |
| L3. Attestation and evidence normalization. [6,7,24,25,30,33,34,86] | Attester-verifier roles, claim formats, measured state, freshness and bounded implementation results. | A measurement claim requires an appraisal process, reference values and freshness context distinct from identity and lifecycle history. |
| L4. Provenance and lifecycle events. [2,21–24,26,37,38,50,51,68] | Event capture and custody models, content provenance, software supply-chain steps and semantic provenance. | Represent temporal transformations and transfers separately from point-in-time attestations and later log registration. |
| L5. Registry and transparency. [17,19,24,26,32,42,49] | Registration receipts, signed or forward-integrity logs, consistency and inclusion proofs. | Protect the integrity and observability of submitted statements; maintain distinction from whether the originating event was true or complete. |
| L6. Policy, verification and risk decision. [6,8,12,25,31,40,65,88] | Appraisal/reliance relationships, selective disclosure, authorization and regulatory decision context. | An authorized decision applies purpose, policy and status to evidence; these decision criteria differ from evidence storage. |
| L7. Governance, audit and institutional feedback. [5,18,37,41,43,47,62,69] | Institutional safeguards, incident reconstruction, custody accountability and declaration reconciliation. | Audits, institutional corrective action and retrospective reconstruction have a longer scope and different authority than one operational appraisal. |

Source: Authors' claim-level and architectural synthesis of the sources cited in the main manuscript.

## Cross-Layer Failure-Mode Traceability

Table S10 gives the source-to-inference chain for each proposed boundary check. The middle column records the published mechanism property and its explicit evidentiary scope; the final column states the additional check demanded at the receiving function and a possible failure if the check is missing. These are analytical tests of evidence sufficiency across separately documented functions, not a count or prevalence ranking of real incidents. The deterministic V1 and V2 examples in Tables S11-S12 exercise the L4-L5 and L5-L6 checks under policy P3; the CAR case [18]

illustrates retrospective evidence gaps without asserting that a particular cryptographic service was deployed.

Table S10. Source-defined properties and reproducible verification obligations at assurance boundaries.

| **Boundary; reference basis** | **Documented guarantee and explicit scope** | **Downstream verification obligation; analytical implication** |
|---|---|---|
| L1-L2: item to identity [1,3,14,29] | Item identifiers and physical identity mechanisms support observed or enrolled association; environmental and enrollment assumptions qualify physical roots. | Re-observe the item and compare its current identifier with enrollment and custodian records. A matching stored credential alone cannot exclude intervening replacement. |
| L2-L3: identity to attestation [6,7,30,33] | RATS separates device Evidence, Verifier appraisal and Relying Party reliance. Valid evidence depends on target, freshness and appraisal inputs. | Verify evidence/key binding, nonce freshness, endorsements and current reference values. Authenticated identity alone does not establish a compliant measured state. |
| L3-L4: state to event history [2,6,21,22] | Attestation supports a bounded state claim; EPCIS/PROV capture events and relationships across transformations or transfers. | Match the appraisal to the correct entity, operation and event identifier. A fresh quote does not alone demonstrate an accountable custody transition. |
| L4-L5: required event to receipt [2,17,19,32] | Event models describe submitted observations; SCITT receipts attest registration of a particular signed statement, with CT proofs covering inclusion/consistency. | Require a receipt for the expected event and reconcile with independent evidence. V1: signed T9 absent; F4 receipt valid, but K4/K5 fail -> HOLD. |
| L5-L6: receipt to authorization [6,8,31,32] | Registration proves a statement was included; policy and appraisal decide whether stated evidence permits a requested action. | Check current purpose, status, intended recipient and independent custody document. V2: T9 receipt valid; manifest lists C rather than B, so K6 fails -> HOLD. |
| L6-L7: decision to investigation [5,18,37,41] | Custody documentation and field investigations support retrospective tracing; safeguards and chain-of-custody models define other evidence and authority. | Archive policy version, inputs and decision; compare with independent observations. CAR [18] supports end-use review, while technical attestation/log evidence is unreported. |
| Cross-cutting: credential state [6,30,35,36] | Trust-anchor status and credential scope can change over time; federation requires independently maintained authority relations. | Use decision-time credential/status snapshots and revocation propagation checks; stale authority can invalidate reliance even when signatures verify. |

| Boundary; reference basis | Documented guarantee and explicit scope | Downstream verification obligation; analytical implication |
|---|---|---|
| Cross-cutting: privacy [8,20,31,74] | Selective disclosure controls presented claim data; authorization and metadata linkability are separate considerations. | Check purpose and minimum disclosure; account for stable identifiers and repeated queries. A selective proof does not by itself establish unlinkability. |
| Cross-cutting: AI model release[23–26,68] | MalHug documents malicious code in model-hub artifacts[23]; Atlas records checkpoint lineage and TEE claims [24]; Laminator attests scoped ML properties [25]; model signing and provenance claims are analyzed in [26]. | Proposed release check: match the model digest, publisher credential, attested training/inference context, declared transformation, registered manifest and recipient policy. A valid signature alone cannot establish benign content, training-data completeness or transfer authorization. |

Source: Authors' comparison of source-defined assurance scopes and downstream decision requirements. Cited properties are documented in the listed references; interface tests and policy responses are the authors' synthesis. V1 and V2 results are specified in Tables S11-S12.

## Applied Case Analysis and Reproducible Decision Trace

The two custody-oriented applications of the reference architecture use different evidence modes. For the retrospective CAR case [18], "Observed" denotes a fact recorded in the cited investigation, "Discrepancy" denotes conflicting reported civil-intent and military-use evidence, "Partial" denotes an evidenced trade relationship without unit-level cryptographic assurance, and "Unknown" denotes a verification function for which the public account supplies no usable record. "Escalate" is a proposed analytical disposition, not an action reported by CAR. Table 6 of the manuscript records the seven checks and the evidence available for each.

The deterministic replay is a fully specified logical example, separate from the CAR investigation. Unit M-017, custodians A/B/C, event identifiers T9/F4, digest H2 and policy P3 are synthetic labels. All variants stipulate that observation and enrolled identifier match, the credential is valid at decision time, the attestation signature and challenge-response are valid and fresh, the measured firmware digest is the approved H2, and the intended end use is authorized. Receipt verification is given as a Boolean input for the specified statement: a valid F4 receipt does not prove inclusion of transfer T9. The replay evaluates logical decisions under these stated assumptions.

Table S11. Complete input differences for the three synthetic decision variants.

| Input field | V0 / complete | V1 / missing | V2 / conflict |
|---|---|---|---|
| Observed/enrolled unit | M-017 / M-017 | M-017 / M-017 | M-017 / M-017 |
| Credential status, target | valid; M-017 | valid; M-017 | valid; M-017 |

| Input field | V0 / complete | V1 / missing | V2 / conflict |
|---|---|---|---|
| Signed fresh attestation | valid; digest H2 | valid; digest H2 | valid; digest H2 |
| Approved image/reference | H2; current | H2; current | H2; current |
| Requested transfer | A -> B | A -> B | A -> B |
| Signed transfer T9 | T9; A -> B; valid | absent | T9; A -> B; valid |
| Event registry receipt | verified for T9 | verified for F4 only | verified for T9 |
| Independent manifest | B | B | C |
| Policy, allowed end use | P3; yes | P3; yes | P3; yes |
| Audit package | observation, appraisal, transfer, receipt, manifest, decision | observation, appraisal, missing-event marker, F4 receipt, manifest, hold | observation, appraisal, transfer, T9 receipt, conflict, hold |

Decision rule P3 (evaluated in order): K1 checks the equality of observed and enrolled unit; K2 requires a valid credential bound to that unit; K3 requires a valid fresh attestation whose measured digest equals an approved current reference; K4 requires a valid signed transfer event for the requested A-to-B hand-off; K5 requires a verified registry receipt specifically bound to that transfer event; K6 requires an allowed end use and equality of intended recipient, signed-event recipient and independent manifest recipient; K7 archives the evidence identifiers, the policy version, the outcome and any failed checks. Release is authorized if and only if K1 through K6 all pass. Failed or missing checks yield HOLD, followed by auditable escalation through K7. The enumeration K1-K7 corresponds to L1-L7, respectively.

Table S12. Deterministic validation trace and resulting dispositions.

| Check | Declared evaluation | V0 complete | V1 missing | V2 conflict |
|---|---|---|---|---|
| L1 / K1 | Observed unit = enrolled unit | PASS | PASS | PASS |
| L2 / K2 | Valid identity credential, bound to M-017 | PASS | PASS | PASS |
| L3 / K3 | Fresh valid attestation, approved H2 | PASS | PASS | PASS |
| L4 / K4 | Signed event T9 covering A -> B | PASS | FAIL: T9 absent | PASS |
| L5 / K5 | Receipt verified specifically for T9 | PASS | FAIL: only F4 | PASS |
| L6 / K6 | Allowed use and manifest = event recipient = B | PASS | NOT EVALUATED: K4/K5 failed | FAIL: manifest C |
| L7 / K7 | Record evidence, outcome and exception | PASS | PASS: HOLD logged | PASS: conflict logged |
| Final disposition | AUTHORIZE if and only if K1-K6 all pass | AUTHORIZE | HOLD | HOLD |

The three outcomes are obtained by applying the stated predicates to Table S11. The observed difference is purely logical: in V1, registration of F4 does not establish inclusion of T9; in V2, the valid T9 receipt does not resolve the conflict between the recorded transfer recipient and the independent manifest. This test illustrates failure localization and reproducible decisions without quantifying field detection accuracy or deployment effectiveness. The staged evaluation in manuscript Table 8 is obtained directly from Table S12: apply K1-K3

in isolation, then K1-K5, and finally the full release predicate set K1-K6. All three variants satisfy K1-K3; V1 first fails K4/K5, so K6 is not evaluated for that variant; V2 passes K1-K5 and fails K6. Passing a subset does not authorize release under P3. The three cases were deliberately specified to expose these differences, and the comparison supplies no statistical estimate of detection accuracy.

A separate AI-model supply-chain application (Application C in manuscript Section 3.5) compares the published evidence objects rather than applying synthetic policy P3. MalHug documents malicious hub artifacts and dataset-loading scripts[23]; Atlas reports attested checkpoint provenance and registered manifests in ML pipelines [24]; Laminator specifies TEE-supported training and inference property claims [25]; model publisher signing and dataset-property statements are analyzed in [26]. The proposed downstream check binds the exact model version, its transformation and evidence context, the relevant signed record, and recipient-specific policy. These checks extend the evidence-to-architecture mapping in Tables S9–S10; no common deployment dataset or policy P3 release outcome is attributed to the cited AI studies.