



QILAHK

SECURITY YOU CAN PROVE

ADVANCED RESEARCH WHITE PAPER

QTA-R03 / 03 OF 05

Cryptographic Evidence Provenance

Research into traceable transformations selective disclosure and verification

Version 1.1 | October 12, 2026

QILAHK Trust Authority

LOCK / KEY / PROOF

qilahk.com | blog.qilahk.com

Table of contents

Cryptographic Evidence Provenance

Abstract	3
Provenance and the assurance boundary	4
Separate five verification questions	4
Define the artifact manifest	4
Limit the claim of cryptographic assurance	4
A proposed provenance packet	5
Preserve transformations as records	5
Disclose an explicit subset	5
Threat model and illustrative review	6
A transformed configuration artifact	6
Adversarial capabilities	6
Recovery and withheld evidence	6
Evaluation of provenance and disclosure	7
Baselines and experimental corpus	7
Research hypotheses	7
Research gates and references	8
Build and review in stages	8
Open research questions	8
References	8

Abstract

Cryptographic Evidence Provenance

A signed artifact can be authentic to a key and still contain an inaccurate claim. A normalized record can be useful and still lose context essential to review. This paper proposes a QILAHK evidence provenance architecture that separates source authenticity, artifact integrity, transformation lineage, semantic appraisal, and authorized disclosure.

The research design connects content-addressed evidence manifests to explicit transformation records and scoped decision references. It explores how an authorized relying party could verify the integrity of disclosed records while understanding which material remains withheld. The design assumes established cryptographic primitives; it does not introduce a new signature scheme or assert a novel cryptographic security proof.

The evaluation question is whether a provenance packet improves detection of substitution, replay, and transformation errors without disclosing unnecessary evidence. The paper specifies threat assumptions, counterexamples, experimental baselines, and measures for correctness and privacy. Selective disclosure is treated as a protocol design problem that requires a concrete implementation and independent review.

Intended audience

Cryptography engineers, evidence architects, and assurance researchers

Research scope

This white paper presents proposed QILAHK research mechanisms, illustrative scenarios, and evaluation protocols. Its hypotheses are intended for prototype testing and independent review.

QTA-R03 | Version 1.1 | October 12, 2026

Provenance and the assurance boundary

Separate five verification questions

An evidence packet should answer who made a statement, whether its bytes changed, how it was derived, whether it supports the scoped claim, and whether the recipient is permitted to see it. These questions require different controls. A valid signature cannot establish the accuracy of a compromised sensor; an access check cannot establish the correctness of a transformation.

The proposed architecture models original artifacts, derived artifacts, collection activities, transformation activities, and responsible actors. W3C PROV-DM supplies a general vocabulary for entities, activities, agents, and derivation [1]. Cryptographic protection and appraisal rules are additional requirements; provenance relationships alone do not provide them.

Define the artifact manifest

A manifest identifies the subject binding, artifact digest, digest algorithm, byte representation, schema version, collector, observation interval, freshness context, and relevant policy references. It declares whether the digest covers original bytes, canonicalized structured data, or both. The verifier must reject ambiguous representations rather than choose a serialization implicitly.

For structured artifacts, canonicalization needs a precise profile, including number representation, Unicode handling, ordering, duplicate-field treatment, and permitted schema extensions. Two implementations should agree on the bytes being signed. Hash agility requires recording the algorithm and a migration procedure rather than treating every digest as interchangeable.

Limit the claim of cryptographic assurance

Content addressing binds a reference to bytes under assumptions about the chosen hash function. It does not make the referenced bytes available, establish their truth, or prove when they were collected. A signing key needs a validated relationship to a collector or authority, and freshness requires a protocol or time model appropriate to the intended use.

A proposed provenance packet

Preserve transformations as records

Each transformation record names the input artifact digests, output digest, procedure version, parameters, execution identity, and time context. A deterministic transformation should be reproducible from these inputs. A nondeterministic transformation must record the relevant execution context and cannot promise bit-for-bit replay unless all required factors are preserved.

The decision reference binds the exact evidence manifest and criteria version reviewed. If a transformation is corrected later, create a new artifact and a new derivation record. Do not replace the old object behind a stable evidence reference. The reviewer needs to know whether a correction changes the basis of an earlier determination.

Disclose an explicit subset

A packet can contain a manifest, disclosed artifacts, disclosure authorization, and verification material. For a committed collection, a Merkle inclusion proof can demonstrate that a disclosed digest belongs to that collection. It does not prove that the collection is complete or that omitted artifacts are harmless. The recipient must know the disclosure scope and what remains unavailable.

RFC 9162 specifies Merkle inclusion and consistency mechanisms for Certificate Transparency [2]. Their structural ideas can inform a research prototype, but using a Merkle tree for evidence does not make the prototype a Certificate Transparency implementation. Leaf construction, domain separation, tree ordering, signed roots, and proof verification must be defined for the new protocol.

Proposed packet components

Component	Verification purpose
Source manifest	Identify subject scope bytes algorithms and collector context
Derivation records	Trace transformations and their input and output versions
Disclosed artifacts	Permit substantive review of authorized evidence
Commitment and proof	Verify membership of disclosed items in a specified collection
Decision binding	Identify the criteria and evidence used by the authority
Disclosure description	Explain omissions and recipient authorization

Commitments to low-entropy sensitive values can allow guessing attacks. Avoid publishing direct digests of predictable personal attributes. A privacy design may need randomized commitments or other specialized mechanisms, plus analysis of metadata leakage and recipient correlation. Such mechanisms require separate cryptographic review.

Threat model and illustrative review

A transformed configuration artifact

An illustrative collector emits a structured device observation. A parser extracts the boot configuration, and a normalization service maps the source value to a control result. The authority reviews that result. The packet should permit an authorized reviewer to identify the original source, inspect the mapping rule, and determine whether the mapping supports the stated property.

If a parser incorrectly treats an unknown source value as enabled, signatures on the source and output can both validate. The derivation chain makes the error traceable but does not prevent it. Independent semantic checks or replay tests are needed. This is the central distinction between cryptographic integrity and appraisal correctness.

Adversarial capabilities

The threat model includes artifact substitution, manifest alteration, replay, mixed-subject packets, unauthorized disclosure, transformation compromise, and signing-key theft. It assumes at least one relevant trust anchor remains uncompromised and the verifier enforces the specified algorithm and key-validation policy. A compromised authorized collector may sign fabricated evidence.

Bind freshness challenges, intended recipient, and assessment context where the protocol requires them. Pin the decision to an artifact collection and policy version. A verifier that checks signatures but ignores audience or subject binding can accept an authentic packet for the wrong purpose.

Recovery and withheld evidence

Key compromise requires a documented invalidation and reassessment procedure. A signature verified today may not resolve whether the key was authorized at the observation time. Retain validated time evidence and key-history context when the assurance model depends on that distinction.

Withholding evidence changes what a recipient can establish. The packet should describe the withheld scope without exposing prohibited details. A recipient may verify integrity while lacking enough information to appraise the claim. Report that limitation rather than present partial cryptographic verification as complete substantive assurance.

Evaluation of provenance and disclosure

Baselines and experimental corpus

Compare unsigned artifact export, signed artifacts without lineage, full provenance packets, and scoped disclosure packets. Use the same collection and transformation workloads. Inject incorrect mappings, reordered fields, duplicate keys, digest substitution, cross-subject reuse, stale challenges, and unavailable source objects.

Test interoperability with two independently implemented serializers and verifiers. A self-consistent implementation can conceal the same canonicalization error on both sides. Separate primitive validation from protocol validation and semantic appraisal, and report which layer detects each injected fault.

Research hypotheses

H1 proposes that provenance packets improve localization of transformation errors relative to signed-artifact export. H2 proposes that scoped packets reduce disclosed sensitive fields while preserving the verification task defined for the recipient. Neither hypothesis implies that all claims remain appraisable after redaction.

Measure false acceptance and false rejection by attack family. Include the inability to retrieve source bytes as a distinct result. For privacy, assess exposed fields, linkable metadata, and inference from low-entropy commitments. Any claimed privacy property should specify the adversary's auxiliary knowledge and the disclosure protocol used.

Research measures

Measure	Definition
Fault detection	Rejected injected-invalid packets divided by all injected-invalid packets
Localization	Cases with the correct faulty transformation identified divided by faulty cases reviewed
Task completion	Recipient reviews completed with the declared scope divided by attempted reviews
Disclosure volume	Sensitive fields and bytes disclosed per recipient task
Verifier cost	Verification time memory and packet size across collection sizes

Do not interpret a low byte count as privacy by itself. A small packet can reveal a stable identifier or a decisive sensitive attribute. Use field-level and linkage analysis alongside volume, and preregister which recipient tasks are expected to remain possible after selective disclosure.

Research gates and references

Build and review in stages

Start with byte-level manifests and schema validation. Add explicit derivation records, deterministic transformation replay, and pinned decision references. Introduce selective disclosure only after defining the recipient's task and the meaning of omissions. Commission protocol and privacy review before public commitments are deployed.

Acceptance gates should require independent canonicalization agreement, rejection of mixed-subject and stale-context packets, reproducible lineage inspection, and visible incomplete-disclosure outcomes. A field that cannot be verified should remain unverified even if the rest of the packet validates.

Open research questions

Which transformations can be independently replayed without disclosing source secrets? What evidence is needed to bind a source key to a physical device? How can key compromise history be incorporated into later reviews? Which disclosure mechanisms preserve useful appraisal while limiting cross-recipient correlation?

The proposed contribution is a verification contract that exposes the boundary between integrity and meaning. Its practical value depends on independently implemented verifiers, adequate collector assurance, and a disclosure policy appropriate to the evidence being shared.

References

[1] W3C PROV DM The PROV Data Model

Provenance concepts and derivation relationships.

<https://www.w3.org/TR/prov-dm/>

[2] IETF RFC 9162 Certificate Transparency Version 2

Merkle proof concepts used as a design reference.

<https://www.rfc-editor.org/rfc/rfc9162>

[3] QILAHK Journal What Makes Security Evidence Trustworthy

Evidence scope integrity and provenance framing.

<https://blog.qilahk.com/what-makes-evidence-trustworthy.html>

Reference editions and source pages checked October 12, 2026. The proposed mechanisms and evaluation plans in this paper are research designs, not reported experimental findings.