VeritasChain Releases CAP-SRP, the First Open Standard to Cryptographically Prove AI Content Was Never Generated

CAP-SRP enables regulators and auditors to verify, with cryptographic evidence, that harmful AI generations were blocked and never occurred.

Logo of the VeritasChain Standards Organization (VSO), a neutral standards body developing cryptographic audit and provenance frameworks for AI systems.

TOKYO, JAPAN, January 10, 2026 /EINPresswire.com/ -- VeritasChain Standards Organization (VSO) today announced the public release of CAP-SRP v0.1.0, an open proof-of-concept specification for cryptographically verifiable AI safe refusal provenance.

CAP-SRP addresses a critical audit gap in modern generative AI systems: while generated content is often logged, refused or blocked generations typically leave no verifiable trace. As a result, regulators and auditors cannot independently confirm that safeguards actually worked.

CAP-SRP introduces a minimal audit primitive that makes non-generation provable.

 The Compliance Problem

Current AI governance practices focus on what systems generate.
However, many regulatory obligations concern content that must not exist at all, such as child sexual abuse material, non-consensual intimate imagery, or other illegal outputs.

When regulators ask, “Can you prove that this content was never generated?”, most AI systems cannot provide a reliable answer. Internal logs, policy documents, and statistical reports are insufficient for independent verification.

This creates a structural compliance risk in the context of emerging regulatory frameworks, including the EU AI Act, the Digital Services Act (DSA), and laws addressing non-consensual and harmful imagery.

> AI systems can log what they generate, but they rarely provide proof of what they refused to generate. CAP-SRP addresses this audit gap by making non-generation cryptographically verifiable.”
>
> *Tokachi Kamimura, VeritasChain Standards Organization*

 What CAP-SRP Provides

CAP-SRP defines an auditable mechanism for recording AI refusal decisions with cryptographic integrity.

The specification demonstrates the following properties:

Every generation attempt is recorded as a verifiable event.
Every refusal decision is cryptographically logged.
Completeness is provable, meaning every attempt has exactly one outcome.
Logs are tamper-evident through hash chaining and digital signatures.
Evidence can be exported in a regulator-ready evidence pack for third-party audit.

Importantly, unsafe prompts and prohibited content are never stored. Only cryptographic hashes are recorded, preserving privacy while enabling verification.

 Negative Proof as an Audit Primitive

Most existing provenance and authenticity standards focus on positive artifacts, such as generated images or documents. They explicitly do not make claims about content that was never created.

CAP-SRP addresses this gap by enabling negative proof: mathematical evidence that a specific harmful generation did not occur.

This capability is essential for regulated AI environments, where the absence of illegal content must be demonstrable, not merely asserted.

 World-First Evidence Assessment

Prior to release, VeritasChain conducted a consolidated prior-art assessment covering more than 250 academic, industry, patent, and standards sources across five independent research engines.

The assessment concluded that CAP-SRP represents the world’s first open specification that combines:

Cryptographic refusal logging
Completeness verification of attempt-to-outcome events

Exportable evidence packs for regulatory and audit use

While related ideas exist in adjacent domains and parallel projects have emerged recently, no open standard currently defines this capability for AI content moderation.

The full World-First Evidence Report is publicly available in the project repository.

 Scope and Limitations

CAP-SRP is not a content moderation product, policy framework, or legal advisory tool. It does not determine what content should be allowed or blocked.

Instead, it provides a narrow but critical building block: verifiable evidence that certain AI actions did not happen.

External anchoring mechanisms, such as blockchain or trusted timestamp authorities, are optional in this proof-of-concept and recommended for production deployments.

 Availability

CAP-SRP v0.1.0 is publicly available under an open license.

Repository:
https://github.com/veritaschain/cap-safe-refusal-provenance

World-First Evidence Report:
https://github.com/veritaschain/cap-safe-refusal-provenance/blob/main/Cap-srp-world-first-evidence-report.md

 About VeritasChain Standards Organization

VeritasChain Standards Organization (VSO) is an independent standards body focused on cryptographically verifiable audit frameworks for AI and algorithmic systems. VSO develops open specifications that enable regulators, auditors, and market participants to move from trust-based to evidence-based oversight.

 Key Message

Verify, Don’t Trust.

TOKACHI KAMIMURA
VeritasChain Co., Ltd.
kamimura@veritaschain.org

Visit us on social media:
LinkedIn
Facebook
YouTube
X
Other
