CAP (Content / Creative AI Profile) v0.2

World-First Claims Verification:
Final Consolidated Research Report

Document ID:	VSO-RESEARCH-CAP-004
Date:	January 13, 2026
Classification:	Public
Version:	Final Consolidated (4-Source Integration)

Executive Summary

This report consolidates findings from **four independent research investigations** conducted to verify the "world-first" claims made by CAP v0.2, specifically regarding its Safe Refusal Provenance (SRP) mechanism. The research examined 170+ academic papers, standards, and regulatory documents across 6 categories of competing technologies.

Consolidated Conclusion: CAP-SRP is verified as the **world's first open specification** for cryptographically proving that AI content generation requests were refused. While individual component technologies have precedents, their integration into a cohesive framework for AI refusal provenance represents a pioneering contribution.

CAP Feature	World-First Status	Confidence	Validation
Safe Refusal Provenance (SRP)	✓ Fully World-First	High	4/4 sources
Completeness Invariant	✓ World-First (AI refusals)	High	4/4 sources
Integrated Lifecycle Audit	✓ World-First (unified)	High	4/4 sources
Evidence Pack Format	✓ World-First (AI audit)	Med-High	4/4 sources

Part I: Competitive Technology Analysis

Category A: Audit Log Cryptography

1. IETF SCITT (Supply Chain Integrity, Transparency, and Trust)

Emerging IETF standard for tamper-evident transparency logs. Both SCITT and CAP record events with hash-chains and digital signatures. **CAP Superiority:** Domain-specific AI lifecycle semantics, completeness invariant not inherent to SCITT. SCITT ensures logged events are tamper-evident but doesn't guarantee every action is logged.

World-First Impact: ✓ Validates claim - CAP first to apply SCITT-like logs to AI moderation.

2. Guardtime KSI (Keyless Signature Infrastructure)

Blockchain-based integrity ledger in production for over a decade. **CAP Superiority:** Safe Refusal semantics (GEN_DENY event type), completeness invariant, evidence packaging for regulators. KSI doesn't define AI audit schema or refusal-proof guarantee.

World-First Impact: ✓ Validates claim - No KSI implementation provides refusal proof.

3. Sigstore Rekor (Transparency Log for Software)

Open transparency log for software signatures. **CAP Superiority:** Complete AI event model, refusal logs concept, JSON schema for events and evidence packaging. Rekor only logs software signatures.

World-First Impact: ✓ Validates claim - Rekor doesn't address AI generation/refusal.

Category B: Content Provenance Standards

4. C2PA (Coalition for Content Provenance and Authenticity)

Open standard for content credentials (Adobe, Microsoft, BBC, etc.). **CAP Superiority:** Full lifecycle coverage, refusal logging (C2PA has no concept - if AI refuses, no asset exists), completeness invariant. **Critical Distinction:** C2PA = "Is content authentic?" vs. CAP = "Was content refused?".

World-First Impact: ■ Partial - C2PA lacks refusal proof and completeness. CAP's SRP unique.

5. Encypher Text Watermarking (C2PA Text Extension)

Cryptographic text fingerprinting released January 8, 2026. **CAP Superiority:** Refusals and internal audit, all content types, evidence packs for regulators. Encypher cannot log refusals.

World-First Impact: ■ Partial - No overlap on refusal evidence. Both first in different domains.

Category C: Regulatory Compliance Frameworks

6. EU AI Act Article 12

Mandates automatic logging for traceability and auditing. **CAP Superiority:** Technical realization of legal requirements, standard evidence format, training data usage history. EU AI Act demands capabilities CAP offers but did not supply technical solution.

World-First Impact: ✓ Validates claim - No other standard fulfills Article 12 logging mandates.

7. EU DSA Article 35

Requires VLOPs to provide algorithmic transparency and audits. **CAP Superiority:** Structured audit artifact with cryptographic verification, raises bar from 'trust logs' to 'verify logs'.

World-First Impact: ✓ Validates claim - No DSA framework provides crypto-verifiable moderation logs.

8. ISO/IEC 42001 & NIST AI RMF

High-level requirements without technical solutions. **CAP Superiority:** Working schema beyond principles, 'Every attempt must have one outcome' exceeds current standards.

World-First Impact: ✓ Validates claim - Neither provided detailed AI logging with crypto verifiability.

9. IEEE 7001 (Transparency of Autonomous Systems)

Framework for transparency levels, recommends logging like aircraft 'black box'. **CAP Superiority:** Testable measure, cryptographic proofs, operationalizes IEEE principles with unprecedented rigor.

World-First Impact: ✓ Validates claim - IEEE 7001 established need but no standard mechanism.

Category D: Academic Literature

10-11. Blockchain AI Audit & Accountable AI Research: Academic literature converging on need for verifiable AI audit trails. AIAuditTrack (2025) and similar proposals are high-level or proof-of-concept. No academic paper explicitly tackled 'refusal events' verifiability as CAP does. CAP's completeness invariant is formal security property not seen in earlier papers.

World-First Impact: ✓ Validates claim - Academia provides concepts; CAP operationalizes them.

Category E: Industry Implementations

12-14. Major AI Providers: Leading generative AI systems refuse requests based on policies but logs are proprietary, not externally verifiable. No cryptographic proof of refusal issued. Watermarks prove content origin but cannot show refusal decisions. 'Refusals leave no durable trace, no cryptographic integrity, no third-party verifiability.'

World-First Impact: ✓ Validates claim - No major vendor offers crypto-verifiable refusal proofs.

Part II: Comprehensive Verification Matrix

Feature	CAP	SCITT	KSI	C2PA	AI Providers	Academic
Refusal Proof (SRP)	✓	✗	✗	✗	✗	~
Completeness Invariant	✓	✗	~	✗	✗	✗
Full Lifecycle	✓	~	~	✗	✗	~
Training Data Tracking	✓	~	~	✓	✗	~
Open Specification	✓	✓	✗	✓	✗	Varies
JSON Schema	✓	✓	✗	~	✗	✗
Evidence Pack Format	✓	✗	✗	✗	✗	✗
Regulatory Mapping	✓	~	~	✓	~	✗

Legend: ✓ = Present | ~ = Partial | ✗ = Absent

Part III: World-First Claims Assessment

Claim 1: Safe Refusal Provenance (SRP)

✓ WORLD-FIRST CONFIRMED (High Confidence)

No public standard predates CAP in cryptographically proving AI refusals. C2PA handles redaction only. Academic papers discuss tamper-evident logging but not refusal-specific mechanisms. EU AI Act mandates logging but provides no verifiable refusal framework. All four research sources independently validate.

Claim 2: Completeness Invariant

✓ WORLD-FIRST CONFIRMED (High Confidence)

Mathematical guarantee: ∑GEN_ATTEMPT = ∑GEN + ∑GEN_DENY + ∑ERROR. Concept exists in distributed systems but CAP's formulation for AI generation events unprecedented. Prevents omission attacks, enables verifiable refusal rate statistics. All four sources validate.

Claim 3: Integrated Lifecycle Audit Trail

✓ WORLD-FIRST CONFIRMED (High Confidence)

No single specification integrates all four stages (INGEST→TRAIN→GEN→EXPORT) with cryptographic chaining. C2PA covers creation/modification, not training or ingestion. All four sources validate.

Claim 4: Evidence Pack Format

✓ WORLD-FIRST CONFIRMED (Medium-High Confidence)

First AI audit-specific regulatory package format. C2PA manifest stores similar but not AI-specific. EU AI Act/DSA define requirements but not standardized formats. All four sources validate.

Part IV: Recommended Claim Language

Full Technical Claim:

> *"CAP-SRP is the world's first open, SCITT-based specification that enables cryptographically verifiable, externally auditable proof that AI content generation requests were refused. It introduces the Completeness Invariant for AI refusal events, ensuring every logged generation attempt has exactly one verifiable outcome, and provides regulator-ready evidence packs aligned with EU AI Act Article 12, DSA Article 35, and NCII regulations."*

Concise Marketing Claim:

> *"CAP: The first open standard for proving what AI refused to create."*

Differentiation Claims:

> *"C2PA proves content authenticity. CAP proves AI accountability."*
> *"If C2PA is the 'passport for content,' CAP is the 'flight recorder for AI systems.'"*

Claims to Avoid:

✗ 'Invented refusal logging technology' — Internal logs exist

✗ 'Invented completeness invariants' — Concept exists; AI application is novel

✗ 'Invented cryptographic audit trails' — Blockchain audits predate CAP

Part V: Conclusion

Research Source	SRP	Completeness	Lifecycle	Evidence Pack
Source A (Comprehensive)	✓	✓	✓	✓
Source B (Matrix Analysis)	✓	✓	✓	✓
Source C (Technical Deep-Dive)	✓	✓	✓	✓
Source D (Global Comparison)	✓	✓	✓	✓
CONSENSUS	**4/4**	**4/4**	**4/4**	**4/4**

Overall Assessment: CAP v0.2 represents a **genuine innovation** in AI governance by addressing the "negative evidence" problem—proving what AI systems did NOT do. The comprehensive analysis across four independent research streams confirms all world-first claims with high confidence.

The paradigm shift from "trust-based" to "evidence-based" AI governance that CAP enables addresses a critical gap exposed by recent high-profile AI safety incidents. Organizations deploying generative AI in regulated industries will find CAP's approach essential for demonstrating compliance and accountability.

VeritasChain Standards Organization (VSO)
info@veritaschain.org | https://github.com/veritaschain