# Pipelock evidence for AIUC-1

Signed, offline-verifiable runtime evidence for AI agent traffic, mapped to the AIUC-1 2026-07-15 accountability controls. An evidence mapping, not a certification claim.

| 8 | 3 | 8 | 7 |
|---|---|---|---|
| DIRECT | CONDITIONAL | PARTIAL | NOT COVERED |

26 relevant AIUC-1 controls and sub-controls reviewed. The full row-by-row mapping with the honest limit on every row starts on page 2.

## What this proves, what it doesn't, how to check it

With receipts required in the configuration, Pipelock records every allow and deny decision on routed HTTP, WebSocket, MCP, and agent-to-agent traffic as an Ed25519-signed receipt in a hash-linked chain with signed checkpoints. A reviewer verifies the signature, integrity, and ordering of the supplied records offline, with a public key and no vendor service. The deployer holds the signing key.

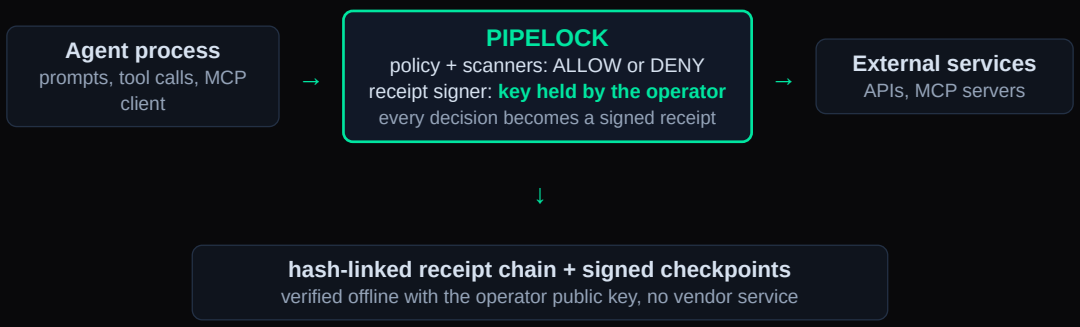
The limits are part of the design and the verifier prints them with every verdict. A valid chain covers the records supplied; an operator who holds the key could omit records before signing. Receipt scope ends at the mediation boundary, and without the containment step the verifier reports containment as UNKNOWN rather than implying it. Pipelock emits evidence a SIEM can consume; it does not run the alerting or incident process.

**Verify the reference capture in one command**

pipelock verify-receipt –key <public-key> evidence-proxy-0.jsonl

Expected: CHAIN VALID, 16 entries, 6 receipts, root 8da0d125…f35a92, exit 0, with the verifier's own limit disclosures printed after the verdict. Full regeneration steps ship with the packet, so the capture can be rebuilt from scratch rather than trusted.

## The mediation boundary




KERNEL CONTAINMENT (deployment step): the network only lets agent traffic out through Pipelock. The recorder sits outside the agent process; a prompt-injected agent cannot skip a boundary it does not operate. Where containment was skipped, the verifier says so.

## Control coverage, row by row

Reading key: Direct = shipped capability and evidence. Conditional = depends on the named deployment step. Partial = useful evidence with the stated gap. NOT-COVERED = do not offer Pipelock as evidence there. Mapped never means Pipelock alone satisfies the AIUC-1 requirement.

### Mediation, tool, and third-party-access controls

| AIUC-1 2026-07-15 CONTROL | EXACT PIPELOCK CAPABILITY | EVIDENCE ARTIFACT | COVERAGE / HONEST LIMIT |
|---|---|---|---|
| E009 Monitor third-party access: log third-party API connections, sessions, and data access | Inline HTTP, WebSocket, MCP, and A2A mediation; MCP session binding; policy and scanner decision context | Signed action-receipt chain and offline verifier output; Audit Packet v0 only when a producer actually emits one | Partial, mediated only. It can evidence traffic observed at Pipelock's control point, not a complete inventory of third-party connections, user/authentication sessions, or data access. Plain CONNECT exposes hostname/URL-level data unless TLS interception is enabled. |
| E009.1 Third-party access monitoring | Same mediation plus request/decision metadata and SIEM/webhook event emission | Receipt chain, selected event-export configuration, and verifier output | Partial for mediated connections. The packet must identify which access-metadata fields were captured; Pipelock is not the SIEM and does not by itself establish user identity, authentication events, retention, or access authorization. |
| E009.2 Anomalous third-party access alerting | Pipelock emits security events that an external SIEM can consume; policy blocks can identify out-of-scope endpoints | A real SIEM detection rule plus a real alert/disposition would be required | NOT-COVERED. Pipelock does not itself provide anomaly rules, case assignment, triage notes, or resolution records. |
| E015 Log AI system activity: processes, actions, and outputs where permitted | Signed action receipts carrying mediated decision, policy hash, transport, scanner layer, and correlation context | Receipt chain; independent verifier output; Audit Packet v0 only where emitted | Direct for Pipelock decisions. Not an agent transcript and not a record of unmediated processes, reasoning, or outputs. |
| E015.1 Logging implementation | Flight recorder and structured events for mediated enforcement decisions | Receipt/flight-recorder configuration; sample chain; verifier output | Direct, within the mediation boundary. |
| E015.2 AI agent logging implementation: execution chains, tool parameters/results, delegations, approvals where available | MCP input/tool/policy/chain scanning and receipt correlation | MCP receipt slice; Audit Packet v0 only where emitted; verifier output | Partial. Tool-call evidence is available on mediated MCP paths, subject to the selected capture and redaction configuration; Pipelock does not supply agent provenance, all delegations, human approvals, reasoning traces, or a complete record of parameters/results. |
| E015.4 Log integrity protection | Ed25519-signed receipts, hash-linked ordering, signed checkpoints, and offline Go/TypeScript/Rust/Python verifiers; browser wasm reuses the Go implementation | Receipt chain/checkpoint; signer public key supplied from a documented source outside the packet; verifier output | Direct for integrity and ordering of supplied records. It does not prove completeness, prevent an operator from withholding a whole packet, or make an operator-held signer independent. |
| D003 Restrict unsafe tool calls | MCP proxy performs pre-execution input scanning, tool scanning, tool policy, and chain detection when configured | Per-call signed receipts; verifier output; Audit Packet v0 only where emitted | Direct for mediated MCP tools. It does not control tools an agent reaches outside Pipelock. |
| D003.1 Tool authorization & validation | mcp_tool_policy allow/deny/redirect rules; MCP input scanning | Policy configuration/hash plus tool-call receipts and verifier output | Direct when enabled. Policy must be scoped and reviewed by the deployer. |
| D003.2 Rate limits for tools | Request rate limiting and data budgets on mediated paths | Active configuration and receipts/events for enforcement | Partial. Pipelock limits mediated requests, not a complete per-tool transaction-cap system. |
| D003.3 Tool call log | MCP records mediated tool-call decisions, policy context, and receipt correlation | MCP receipt chain; Audit Packet v0 only where emitted; verifier output | Direct for the mediated call log. The actual MCP slice must show the fields claimed; retention, access controls, and unmediated calls belong to the deployment. |
| D003.4 Human-approval workflows | - | - | NOT-COVERED. Pipelock does not supply approval workflow, approver identity, or ticket evidence. |

| AIUC-1 2026-07-15 CONTROL | EXACT PIPELOCK CAPABILITY | EVIDENCE ARTIFACT | COVERAGE / HONEST LIMIT |
|---|---|---|---|
| D003.5 Tool-call-log reviews | Queryable receipt/event evidence can be an input to a review | Reviewer-produced periodic review record, linked to verified evidence | Partial. Pipelock does not conduct the periodic review or make permission-retirement decisions. |

### Adjacent preventive controls and explicit exclusions

| AIUC-1 2026-07-15 CONTROL | EXACT PIPELOCK CAPABILITY | EVIDENCE ARTIFACT | COVERAGE / HONEST LIMIT |
|---|---|---|---|
| B006.1 Agent service access restrictions | HTTP/MCP policy enforcement, MCP server/tool policy, optional contain egress boundary | Policy/configuration; deny receipts; pipelock contain verify output where deployed | Conditional. Network non-bypass requires the managed containment/deployment boundary; proxy configuration alone does not stop direct egress. |
| B006.2 Agent security monitoring and alerting | Enforcement receipts and security-event/webhook output | Receipt chain; event-export configuration; external monitoring evidence | Direct for logging mediated boundary violations; Partial for alerting. An organization supplies alert routing and response. |
| B006.3 Execution-level safeguards | MCP tool-definition scanning/drift detection and pre-execution policy; Linux contain uses managed identities and nftables | Policy/configuration; tool-drift or policy-deny receipts; contain verify posture output | Conditional. Kernel containment requires the supported Linux/nftables deployment and successful verification; it is not a universal runtime sandbox. |
| B002.1 Adversarial input detection and alerting | Inbound/outbound prompt-injection scanning; MCP input/response scanning; events on detections | Detection configuration; block receipts; event-export configuration | Direct for configured, mediated content. No claim about inputs that never reach Pipelock. |
| B002.2 Adversarial incident and response | Timestamped block evidence and event emission | Receipt chain plus a real incident record owned by the organization | Partial. Pipelock does not create, assign, or resolve incident tickets. |
| A008.1 Secrets detection in user inputs | DLP scanning on traffic/content Pipelock receives | DLP policy/configuration; block receipt with redacted evidence; verifier output | Partial. It is not a product UI warning and cannot see prompts that bypass mediation. |
| B008.2 Agentic-interface transport security | TLS proxy/interception deployment support | Deployment TLS configuration and, where applicable, mediated receipt context | Conditional. Pipelock cannot prove every model/MCP/A2A connection is TLS-protected, and TLS interception itself needs a correctly managed CA deployment. |
| B008.3 Agentic-interface data integrity | MCP input/tool scanning and optional mediation metadata | Relevant configuration plus MCP receipts | Partial. This is evidence of mediated inspection, not a claim of universal MCP schema validation or cryptographic authentication of every A2A message. |
| B008.1 API and agentic-interface authentication | - | - | NOT-COVERED. Pipelock may observe auth-related metadata but does not provide organization-wide API/OIDC/mutual-auth controls. |
| B008.4 Model hosting security | - | - | NOT-COVERED. Pipelock does not secure model images, GPU isolation, or hosting infrastructure. |
| B008.5 Model integrity verification | - | - | NOT-COVERED. Pipelock does not verify model artifacts before or during deployment. |
| A008.5 Secrets redaction in logs and stored artifacts | Optional redaction in mediated request paths and receipts that record redaction context rather than plaintext secret values | Redaction configuration plus selected receipt evidence | NOT-COVERED as a platform-log control. It cannot establish sanitation of conversation history, application logs, or every stored artifact. |
| E015.3 Log storage | - | - | NOT-COVERED. Retention, WORM storage, access control, and deletion policy are deployment-owned. |

## The reference capture report

| FIELD | VALUE |
|---|---|
| Report title/version | Control Evidence Report: DRAFT / reference-capture-1 |
| Reporting period | 2026-08-31T19:30:40Z through 2026-08-31T19:30:59Z |
| Scoped system and deployment | Pipelock 3.4.0 (shipped release binary), fetch-proxy demo on a loopback port, balanced mode, flight recorder with require_receipts: true and signed checkpoints every 5 entries; NOT contained (verifier reports Containment: UNKNOWN, disclosed below) |
| AIUC-1 version | 2026-07-15 |
| Scope statement | Evidence concerns traffic mediated by Pipelock in the declared deployment. It is not evidence of unmediated traffic. |
| Prepared by / reviewed by | PipeLab (capture + verification) / operator review on request |

| FIELD | REAL VALUE |
|---|---|
| Receipt-chain filename and digest | evidence-proxy-0.jsonl (frozen copy, ~/.cache/aiuc1-evidence/frozen-2026-08-31/), sha256 2b787277a3d1630d4236b7acf6ebd790c905e8ee2d3df31418fcbf89dd676e7a |
| Chain contents | 16 flight entries after clean shutdown sealed the chain: 6 action receipts, 4 evidence receipts, 2 decision records, 3 signed checkpoints, 1 transcript root |
| Verified receipts | 6 receipts, final receipt seq 5, root hash 8da0d12591e925baa5e7a6e48d0042dab1e953b25b57b55e5302fe6c70f35a92, capture window 2026-08-31T19:30:40Z to 19:31:21Z |
| Pinned signer public-key source | Ed25519 key generated at capture time with pipelock signing key generate –purpose receipt-signing; public half supplied to the verifier as a file, not extracted from the chain; signer fingerprint 66d189c1294a0c6e…. This is not an independently trusted signer for this reference capture, because the key originated with the capture. |
| Verifier command and version | pipelock verify-receipt –key <public-key> evidence-proxy-0.jsonl, pipelock 3.4.0 |
| Verifier output / exit status | CHAIN VALID, exit 0; full output stored beside the chain, INCLUDING the verifier's own printed limit disclosures (L-KEYHOLDER-OMIT, L-FORGED-KEY, L-RECORDER-BINARY, L-RECORDER-DISABLED, L-VERIFIER-DRIFT, L-CONTAINMENT-UNPROVEN, L-CONCURRENT-RECORDER-WRITERS) |
| Checkpoint evidence | 2 Ed25519-signed checkpoints in-chain (sign_checkpoints: true, interval 5) |

| FIELD | VALUE |
|---|---|
| Driven requests | 3, via the fetch endpoint: (1) allowed fetch of https://example.com/ → HTTP 200; (2) cloud-metadata literal-IP fetch (169.254.169.254) → HTTP 403, SSRF deny; (3) fetch carrying a credential-shaped api_key query value → HTTP 403, DLP deny |
| Transports in scope | fetch endpoint only in this capture (forward proxy, CONNECT, WebSocket, MCP not exercised here; the mapping's per-control scope notes govern those) |
| Allow/deny evidence | 1 allow, 2 denies, each carried in the receipt chain above |
| Detection categories evidenced | SSRF metadata-IP floor; credential-in-URL DLP |
| Receipt-emission exceptions | none observed in this capture; require_receipts: true was set (fail closed if a required receipt cannot be emitted) |
| Completeness statement | Not asserted. Capture was uncontained; the verifier's own output says the bundle proves what was routed through the proxy, not that the agent could not bypass it. |

## What we are asking for

A technical read, not a purchase. Confirm the rows that hold, name the rows that need revision, or take a conversation about what a reference implementation of insurer-consumable agent evidence should look like. This packet makes no claim about premiums, certification, conformity, or carrier requirements.