ODA3 INSTITUTE

EVIDENCE REQUEST LIST

# PAI-SF™ Evidence Request List

Physical AI Security Framework v1.0 — Assessment Support Artifact

ODA3-2026-07-ERL-PAISF-018 | Public-Facing — Assessment Support Artifact

© ODA3 Pvt Ltd

### Document Control

| | |
|---|---|
| **Doc ID** | ODA3-2026-07-ERL-PAISF-018 |
| **Doc Type** | Evidence Request List |
| **Framework Tag** | PAI-SF™ v1.0 |
| **Audience** | Assessment requester, system owner, security lead, AI governance lead, assurance team |
| **Companion Document** | ODA3-2026-07-NOR-PAISF-005; ODA3-2026-07-CAT-PAISF-006; ODA3-2026-07-AIQ-PAISF-017 |
| **Governing Licence** | GAISSF Ecosystem Licence (GEL v1.0) |
| **Access Classification** | Open / Public assessment support; scoring, certification, and auditor procedures remain Controlled |
| **Publication State** | Publication Candidate — pending formal publish action |
| **Publish Date / Review Cycle** | 01 August 2026 (content finalized); publication date to be confirmed |
| **Classification** | Public-Facing — Assessment Support Artifact |

# PAI-SF™ Evidence Request List

### Physical AI Security Framework v1.0 — Assessment Support Artifact

## Purpose

This evidence request list identifies the starter artifacts normally needed to scope a PAI-SF™ assessment for physically embodied AI systems. It supports evidence collection and assessment planning only; it does not disclose scoring logic, maturity thresholds, auditor procedures, certification decisions, or controlled test protocols.

The evidence requested throughout this list exists to support one underlying question, consistent with **Kinetic Zero Trust** — PAI-SF's core assurance principle: can the requester demonstrate that AI-generated physical actions are validated against independent safety boundaries, rather than trusted on model confidence alone? Section 5's red-flag screen is built around evidence gaps that would leave that question unanswered.

## 1. Submission Principles

| Principle | Expectation |
|---|---|
| Minimum necessary evidence | Provide enough information to establish scope, ownership, physical consequence path, and control-evidence availability. Do not submit secrets or regulated data unless a secure review channel has been agreed. |
| System-specific evidence | Evidence should describe the actual system, deployment, environment, supplier chain, update path, and operating boundary under review rather than generic policy statements alone. |
| Current version basis | Identify document dates, system versions, model or firmware versions, deployment stage, and whether |

| Principle | Expectation |
|---|---|
| | submitted evidence reflects production, pilot, laboratory, or pre-deployment conditions. |
| Boundary clarity | Separate confirmed evidence from assumptions, planned controls, future remediation, inherited platform claims, and vendor assertions not independently verified by the requester. |

## 2. Core Evidence Request

Collect these artifacts before domain-level evidence review. They establish the system boundary, physical consequence path, accountability model, and evidence-handling posture.

| Evidence Area | Requested Evidence |
|---|---|
| System identity and scope | Architecture diagram; system inventory; AI components; sensors; actuators; edge/cloud dependencies; operating sites; deployment stage; accountable owners; user/operator roles. |
| Physical consequence path | Description of how sensing, model inference, autonomy, decision support, or command logic can affect movement, force, access, environment, process state, infrastructure state, payload, vehicle behavior, or human safety. |
| Operating boundary | Operating design domain, site boundary, geography, route, airspace, cell, room, asset zone, public-space exposure, environmental assumptions, and boundary-change approval process. |
| Command and override path | Command flow; authorization controls; timing/replay protections; interlocks; emergency stop; local/remote override; operator handoff; reset and return-to-service rules. |
| Safety and degraded operation | Safe-state definitions; reduced-risk/degraded modes; manual mode; fallback logic; loss-of-perception behavior; communications failure behavior; validation records. |
| Telemetry and incident readiness | Monitoring map; logs for sensor state, autonomy mode, commands, overrides, interlocks, degraded states, incidents, evidence custody, retention, and physical recovery procedure. |
| Supplier and update exposure | Supplier list; integrator responsibilities; firmware/model/ruleset/map updates; staged rollout; rollback; regression testing; maintenance windows; third-party dependencies. |

## 3. Domain-Level Evidence Requests

The table below links requested evidence to the 12 canonical PAI-SF domains. It is a request list, not a scoring rubric, maturity model, or clause-level compliance matrix.

| No. | PAI-SF Domain | Evidence Normally Requested |
|---|---|---|
| 1 | Sensor and Perception Integrity | Sensor inventory, calibration records, perception validation, drift/disagreement handling, degraded-condition tests, spoofing/occlusion assumptions, and known blind spots. |
| 2 | Actuator and Kinetic Command Safety | Physical command architecture, bounded command rules, force/speed/torque or process-state constraints, interlocks, timing/replay controls, and actuator divergence monitoring. |
| 3 | Edge Hardware and Model Attestation | Edge-device inventory, firmware/image provenance, model package identity, attestation evidence, key custody, local runtime trust, and field-device update records. |
| 4 | Autonomy Boundaries | Authorized tasks, autonomy modes, operating-domain definition, boundary detection, escalation criteria, boundary-change approvals, and unauthorized expansion controls. |
| 5 | Safe-State and Degraded-Mode Behavior | Safe-state design, degraded-mode triggers, fallback procedures, independence from compromised compute where claimed, manual-mode transition, and recovery validation. |
| 6 | Human Override and Intervention | Operator authority model, emergency stop/override evidence, handoff procedure, training records, intervention logs, reset criteria, and accountability after intervention. |
| 7 | Runtime Monitoring and Telemetry | Telemetry coverage map, tamper-evidence controls, sensor-confidence logs, autonomy-state logs, command-path logs, override records, safety-interface events, and retention policy. |
| 8 | Incident Response and Physical Recovery | Incident playbooks, physical containment procedure, evidence preservation process, recovery checklist, return-to-service validation, and post-incident learning records. |
| 9 | Supply Chain and Update Assurance | Vendor/integrator inventory, component provenance, firmware/model/ruleset update process, rollout/rollback evidence, supplier attestations, and regression-test results. |
| 10 | Simulation, Testing, and Validation | Scenario library summary, |

| No. | PAI-SF Domain | Evidence Normally Requested |
|---|---|---|
| | | hardware-in-the-loop or simulation evidence, physical-world test records, adversarial/degraded tests, acceptance criteria, and sim-to-real limitations. |
| 11 | Functional Safety Interface | Safety-case linkages, unresolved AI-security gap register, exception escalation route, safety-engineering owner, and evidence showing how security findings reach safety decision-makers. |
| 12 | Physical Operating Environment | Environment description, site maps or operating-zone definitions, lighting/weather/floor/road/airspace/ occupancy assumptions, maintenance zones, obstacles, and environment-change handling. |

## 4. Artifact Handling Notes

| Handling Topic | Instruction |
|---|---|
| Acceptable examples | Diagrams, inventories, policies, playbooks, test summaries, logs, issue registers, supplier attestations, release records, training records, safety-case extracts, and version-controlled design evidence. |
| Prefer excerpts when possible | Use excerpts, redacted copies, summaries, screenshots, or evidence indexes where full artifacts contain secrets, regulated data, customer data, personal data, source code, or sensitive operational detail. |
| Label confidence | Mark evidence as implemented, tested, partially implemented, planned, vendor-asserted, inherited, obsolete, or unknown. This helps separate operating fact from roadmap intention. |
| Preserve chain of custody | For incident, telemetry, or test evidence, preserve timestamps, system versions, collector identity, retention period, and any transformation applied before submission. |

## 5. Intake Red Flags

Flag these conditions before accepting a routine evidence plan. A flag does not establish failure; it means the assessment plan may require deeper scoping, secure evidence handling, or specialist review.

| Flag | Why It Matters |
|---|---|
| Evidence depends only on vendor claims | Requester cannot independently show architecture, update, telemetry, safety-interface, or physical command controls for the deployed system. |

| Flag | Why It Matters |
|---|---|
| Physical command path is unclear | The organization cannot explain how AI behavior can or cannot influence physical effects, or where command bounds and overrides sit. |
| Safe-state evidence is missing | There is no tested response for perception loss, model drift, communications loss, operator loss, command compromise, or degraded physical environment. |
| Telemetry cannot reconstruct incidents | Logs do not preserve enough information about sensor state, autonomy state, command path, override, degraded mode, and physical effect. |
| Updates can alter physical behavior silently | Firmware, model, map, rule, or controller updates can change behavior without staged rollout, rollback, approval, and regression evidence. |

## 6. Notably Absent

| Absent Claim | Clarification |
|---|---|
| No scoring or maturity decision | This evidence request list does not publish scoring thresholds, maturity levels, weighting, pass/fail criteria, auditor procedures, scenario libraries, or certification-decision logic. |
| No certification or compliance claim | Submitting evidence does not establish PAI-SF certification, regulatory compliance, safety certification, medical-device approval, airworthiness, road authorization, building-code compliance, or infrastructure approval. |
| No request for unnecessary secrets | The list is not a request to disclose source code, trade secrets, credentials, regulated data, personal data, or sensitive operational records unless secure handling has been agreed. |
| No substitute for assessment judgment | Evidence availability supports scoping and planning only; relevance, sufficiency, quality, and control performance require system-specific review. |

**On GAISSF Domain 9:** GAISSF Domain 9 control titles are cited via the verified crosswalk (ODA3-2026-07-CSX-PAISF-004), not reproduced independently in this artifact.