**Product Requirements Document (PRD)**

**Project:** Edge-Agent Core (ACAP for agentic AI on Axis hardware)

**Version:** v1.0 Draft

**Date:** 2025-10-19

**Owner:** [Product–Lead Name]

**Status:** Draft – for review

**1. Overview & Strategic Fit**

**Problem Statement:**

Security operations are overloaded by false alarms (animals, lighting changes, foliage, weather) and lack meaningful prioritization. Traditional rule-based video analytics are rigid, high-false-positive, and cloud-dependent. The result: delayed responses, wasted resources, frustrated integrators and end-users.

**Opportunity / Vision:**

Embed an intelligent, on-device agent (“Edge-Agent Core”) on Axis sensors that *verifies, filters, and prioritizes* alarms in real time—edge-first, vendor-agnostic, continuously learning. This becomes the heart of a next-gen security stack, reducing false alarms from >90% to <10%, enabling new recurring revenue models, and positioning Axis as the platform leader in agentic edge AI.

**Strategic Fit:**

* Aligns with Axis’s hardware+software move, transforming cameras into smart endpoints.
* Addresses big pain point in physical security: nuisance alarms + integrator complexity.
* Creates recurring software revenue (per camera license) on top of hardware.
* Sets the stage for broader agentic features (multi-sensor orchestration, predictive threat detection, universal translator).

**2. Objectives & Success Metrics**

**Objectives:**

* Release MVP ACAP “Edge-Agent Core” on Axis OS v12 with on-camera verification of alarms.
* Demonstrate real-world reduction of false alarms by ≥ 80% in pilot sites within 6 months.
* Achieve per‐camera latency ≤ 150 ms for decision loop in typical settings.
* Launch with a “zero-friction” commissioning flow (≤10 minutes setup).
* Build foundation for integrator ecosystem via universal translator and feedback loop.

**Success Metrics (KPIs):**

* False alarm dismissal rate: drop from baseline (e.g., 90%) to < 20% in pilot.
* Genuine threat detection rate: ≥ 95%.
* Decision latency: median ≤ 150 ms; 95th percentile ≤ 300 ms (on target hardware).
* Commissioning time per camera: average ≤ 10 minutes.
* Feedback label participation rate: ≥ 5 operator feedbacks per camera per week in pilot.
* Number of integrators on-boarded: ≥ 10 in first 90 days.
* Monthly recurring revenue (MRR) per camera: tiered (e.g., $5/licence/month) with at least 1000 cams enrolled in year one.

**3. Target Users & Personas**

**Primary Personas:**

* *Security Operations Manager (“SOM”):* Oversees multiple cameras, seeks to reduce wasted dispatches, wants alarms to be meaningful and prioritized.
* *Integrator/Installer (“INT”):* Deploys multi-vendor systems, frustrated by complexity, wants plug-and-play module that “just works” and adds value.
* *Field Technician (“FT”):* Configures cameras on site, needs simple UI, minimal setup, fast verification of correct behavior.

**Use Cases:**

* Camera triggers motion alarm → Edge Agent filters out animal/lighting flicker → no dispatch event, SOM sees “dismissed – low severity” log.
* Camera triggers after-hours door approach → Edge Agent detects approach vector + dwell → severity=high → event forwarded to VMS + mobile alert.
* Multi‐sensor site: radar+camera fusion track enters restricted zone → agent correlates sensor + video → escalates to centralized operator with bounding box and reason.
* Operator corrects a “true false alarm” via mobile app → feedback fed into edge cache → future similar event suppressed more aggressively.

**4. Scope and Boundaries**

**Included in v1 MVP (“Edge-Agent Core”):**

* On-device DL inference (object classes: human, vehicle) + tracker + temporal head.
* Rule engine (zones, schedules, dwell time, approach vector).
* Event output: decision (dismiss / verify / escalate), severity score, reason code, bounding box overlay via RTSP analytics.
* Feedback endpoint on camera for operator labels.
* Commissioning UI for zones/schedules via camera web interface.
* License key activation per device.
* MQTT/webhook output for cloud integration (optional).
* Basic documentation + installer guide + demo kit (False Alarm Gauntlet).

**Out of scope for v1 (but planned for later releases):**

* Long-horizon prediction (>30 s) across complex behavior scenes.
* Extensive third-party sensor integration (beyond optional MQTT ingestion).
* On-device federated learning or full model retraining (feedback loop only threshold/head update).
* Natural language UI or LLM on-camera.
* Full multi-vendor universal translator (planned v1.2).
* Extensive mobile app for operators (beyond simple corrective label UI).

**5. Functional Requirements**

**5.1 Perception & Analytics**

* Camera shall stream video frames via VDO API; infer object classes (human, vehicle) via Larod.
* Camera shall maintain tracker (≤ N tracks) with metadata: class, bounding box, speed, direction.
* Temporal head shall compute dwell time, approach vectors, path to restricted zones.

**5.2 Rule Engine**

* Configurable zones (polygon), schedules (time windows), class allowlists/denylists.
* Configurable thresholds: min\_size\_px, min\_dwell\_ms, approach\_dirs.
* Severity weighting: alpha\*(model\_confidence) + beta\*(sensor\_context) + gamma\*(rule\_score). (Default weights settable.)

**5.3 Decision & Outputs**

* On each event, produce: severity\_score (0-100), reason\_code (predefined set), track\_id, camera\_serial.
* Output via:
  + ONVIF/Event API: state “Threat.Active” and data payload with severity, reason.
  + RTSP analytics overlay: bounding box + reason text.
  + MQTT/Webhook (optional): JSON message to external broker including severity, bbox, track\_id, timestamp.
* Latency from frame capture to decision < 150 ms (median) under standard load.

**5.4 Feedback Loop**

* HTTP(S) endpoint /agent/feedback: accepts POST {camera\_serial, track\_id, correct: boolean, class\_label}.
* Feedback stored locally; at scheduled interval (e.g., weekly) thresholds are updated; optional upload of anonymized stats for federated update.

**5.5 Commissioning & Configuration UI**

* Web UI: step-by-step commissioning wizard: zone drawing, schedule entry, baseline capture.
* Parameter API: read/write JSON for all config items (zones, schedules, weights).
* License Key API: camera reads activation key via VAPIX, enables full features.

**5.6 Non-Functional Requirements**

* Performance: model inference latency ≤ 40 ms, tracker update ≤ 10 ms, event output ≤ 5 ms for ≤ 24 tracks at 10 fps on ARTPEC-8.
* Reliability: System shall operate 24/7 with ≥ 99.5% uptime (excluding maintenance).
* Security: All external communication (MQTT, webhook) uses TLS 1.2+, camera supports secure boot and signed firmware updates.
* Scalability: ACAP shall support up to 5000 cameras per site when centralized management via broker.
* Usability: Commissioning time per camera ≤ 10 minutes for trained technician.

**6. Technical Architecture**

* **Hardware platform:** Axis cameras with ARTPEC-6/7/8 chips; DLPU present (when available); fallback to CPU.
* **Software stack:** Axis OS v12; ACAP v12 Native SDK; compiled for ARMv7hf and AArch64 architectures.
* **Modules:**
  + perception\_module: VDO + Larod inference
  + tracker\_module: lightweight multi‐object tracker
  + temporal\_head: dwell/approach computation
  + fusion\_engine: score fusion & rule engine
  + decision\_module: event emitter & overlay
  + feedback\_handler: web endpoint + local cache
  + config\_ui: web wizard + parameter API
  + license\_manager: license key API
* **Integration layers:** RTSP analytics overlays; ONVIF/Event API; MQTT/Webhook; Parameter API; Edge storage for feedback.
* **Data flow diagram:**

VDO frames → perception\_module → tracker\_module → temporal\_head

\ |

→ fusion\_engine ← sensor metadata (optional)

→ decision\_module → outputs

* **Deployment:** .eap package delivered via Axis Device Manager; subscription license per camera; continuous updates via Axis store.

**7. Acceptance Criteria**

* Pilot site config (≥ 10 cameras) shows false alarm rate reduction ≥ 80% within first month.
* Commissioning wizard validated: technician sets up 5 sample cameras in ≤ 10 minutes each.
* Performance tests: camera with 24 active tracks at 10 fps, latency median ≤ 150 ms, 95th percentile ≤ 300 ms.
* Feedback loop engaged: at least one corrective label per camera in pilot week; subsequent events reflect threshold adjustment.
* Integration tests: RTSP overlay displays correct bounding box + reason; ONVIF event fires correctly; MQTT message arrives at broker with required fields.
* Security audit: external communication encrypted; unauthorized camera access blocked; license key enforceable.

**8. Dependencies & Constraints**

**Dependencies:**

* Axis OS v12 availability on target camera models.
* Larod / DLPU driver support for chosen model architectures.
* Integrator cooperative pilots for initial deployment and feedback.
* Cloud infrastructure (broker/MQTT) for optional telemetry.
* VMS compatibility for RTSP analytics overlays.

**Constraints:**

* On-camera compute & memory budgets: must fit within existing Axis hardware without degrading core video streaming.
* Network bandwidth: must not saturate network (keep overlays and MQTT minimal).
* GDPR/privacy: feedback data and telemetry must be anonymised for federated usage.
* Licensing/regulatory: some installations may require export control compliance for embedded AI.

**9. Risks & Mitigations**

| **Risk** | **Impact** | **Mitigation** |
| --- | --- | --- |
| On-device inference latency too high for target hardware | High | Benchmark early per hardware family; build fallback to CPU with reduced features; allow model compression. |
| False negatives in verification lead to missed threats | Critical | Require pilot evaluation with live/historic data; maintain human-in-loop override and alerts for unknown cases. |
| Integrator resistance due to setup complexity | Moderate | Build intuitive commissioning wizard; provide pre-defined templates; launch “Lite” version free for 5 cams. |
| Security vulnerabilities in feedback/telemetry path | High | Follow Axis security guidelines; external audit; encrypt all communication; no PII stored. |
| Multi-vendor translation layer complexity underestimated | Moderate | Treat translator as v1.2; focus MVP on Axis + ONVIF; build flexible plugin architecture. |

**10. Roadmap & Release Plan**

* **v1.0 (MVP) – “Edge-Agent Core”**: 0-90 days: release ACAP package, pilot deployments (10 sites), feedback loop enabled.
* **v1.2 – “Universal Translator + Short-Horizon Prediction”**: 90-180 days: integrate multi-vendor protocol translator, prediction module (3-10 s), live demo kit.
* **v2.0 – “Predictive Orchestrator”**: 180-365 days: 30-60 s horizon, third-party sensor orchestration, federated model updates, full mobile operator UI.
* **Launch at ISC West/GSX:**
  + Demo booth: False Alarm Gauntlet + Swarm Hand-off
  + Press release: “Axis unleashes first agentic edge camera verifying threats in <200 ms, saving ops millions.”
  + Integrator webinar + early-adopter program launch.

**11. Open Questions & Assumptions**

* Assumes that ARTPEC-8 DLPU can support target DL model plus tracking within latency budget—need hardware benchmarks.
* Assuming integrators will adopt monthly subscription pricing model on top of hardware—they may require sales incentive or bundling.
* How will we handle sites with poor connectivity (for feedback upload/federation)? Offline mode strategy required.
* Need legal review of “outcome-based SLA” (e.g., <10% false alarms or refund) — risk of unbounded liability.
* What is the acceptable range of classes (human/vehicle only?) for MVP to meet performance and latency targets?

**12. Appendices**

* Appendix A: Glossary (e.g., track\_id, severity\_score, reason\_code).
* Appendix B: Proposed reason\_code taxonomy (e.g., “animal”, “lighting\_change”, “authorized\_personnel”, “approach\_vector”, “dwell\_exceed”).
* Appendix C: Model architecture logs and expected benchmarks (to be filled after dev).
* Appendix D: Demo scenario scripts (False Alarm Gauntlet & Swarm Hand-off).
* Appendix E: Licensing terms draft.

**End of Document**