Bounded Inference at the Edge

A compliance architecture for distributed AI inference — gateways and embedded compute — under the EU AI Act, the UK Online Safety Act 2023, and the NIST AI Risk Management Framework

The Novacene Ltd

Version 1.0 · May 2026

About this paper

This is a Novacene-authored brief setting out a compliance architecture for edge AI deployments operating under three converging regulatory regimes: the EU AI Act, the UK Online Safety Act 2023 with its safeguarding satellites, and the NIST AI Risk Management Framework. The architecture spans the full spectrum of edge AI silicon — from premium aggregation gateways at one end, to cost-optimised embedded compute System-on-Modules at the other — and is silicon-agnostic by design. The five layers of the architecture sit unchanged on top of any sufficiently capable edge inference platform. The architecture, the regulatory analysis, and the reference implementations are independent of any specific hardware vendor and are released open under CC BY-NC-SA 4.0. For implementation guidance, or for an audit of an existing edge AI deployment against this architecture, contact The Novacene at legal@thenovacene.com.

Executive summary

Edge AI is now the dominant deployment topology for inference in regulated environments — schools, hospitals, transport networks, public buildings, defence sites. It spans a spectrum, from premium aggregation gateways at one end to cost-optimised embedded compute SOMs at the other. The compliance regimes that govern these settings do not distinguish between the two ends of that spectrum. Three regimes — the EU AI Act, the UK Online Safety Act 2023 with its safeguarding satellites, and the NIST AI Risk Management Framework — are converging on the same procedural requirement: inference must be bounded, auditable, and accountable, with humans demonstrably in the loop.

Most edge AI hardware is currently shipped without a credible accountability layer. This is not a regulatory inconvenience to be lobbied away. It is the fastest-growing line item in technical procurement, and the gap between what vendors ship and what regulated buyers can lawfully operate is widening. The gap is the commercial opportunity.

This paper sets out a five-layer compliance architecture — device, runtime safety middleware, policy-as-code envelope, telemetry, human oversight — that can be implemented on any sufficiently capable edge AI device, gateway or embedded module. Each layer maps to a specific regulatory demand. The pattern is silicon-agnostic by design and is demonstrated, in this paper, across both tiers of edge silicon — premium aggregation gateways at one end, cost-optimised embedded compute SOMs at the other.

We use UK schools as the worked example because it is the hardest deployment in the regulated estate: vulnerable subjects, dense overlapping statutory duties, low tolerance for opacity, and a procurement function that has been lightly scarred by previous AI deployments. If the architecture is defensible there, it is defensible across the regulated estate.

The paper names where existing open-source components — Flare, Verse-Nerves, the verse-ality-agents policy framework — already implement the pattern, and ends with a one-page procurement checklist that buyers and vendors can use to test whether a deployment is operable lawfully on day one.

1. The wedge: why edge AI is the next compliance front

The cloud-AI compliance debate has matured. The edge-AI compliance debate has barely started. That asymmetry is closing fast, and it is closing in the procurement function — not in the model lab.

The deployment shift

Edge AI now runs inference in the very settings where regulators are least tolerant of opacity: schools, custody suites, hospitals, transport hubs, defence installations, public buildings. The shift is not from one device class to another. It is across a spectrum — and any compliance architecture worth the name has to work across that whole spectrum, because the regulatory regimes do.

- **Aggregation gateways.** High-TOPs aggregation of many feeds, typically deployed centrally or per-site. Premium aggregation silicon typically delivers 10–15 Dense TOPs of inference capacity and is offered by all major edge AI silicon vendors. These devices fuse video, audio, and sensor data, and run rich multimodal inference. They are the obvious target for compliance scrutiny because they are visible, expensive, and audit-ready.
- **Embedded compute SOMs.** Cost-optimised modules embedded inside customer products and deployed per-device, often at fleet scale. Cost-optimised embedded compute SOMs are offered by every major edge silicon vendor and a long tail of cellular-IoT module specialists. These devices run lighter, often single-modality inference. They are the part of the deployment that procurement most often forgets is regulated.

The compliance regimes do not distinguish between the two. A model running on a £-thousands aggregation gateway and a model running on a £-tens embedded SOM, in the same school, fall under the same statutory duties. The architecture in Section 3 must therefore be silicon-agnostic and tier-agnostic — and the worked example in this paper demonstrates the same compliance architecture sitting unchanged on either tier.

The three accelerants

Three regulatory developments are arriving on the same desk at roughly the same time.

- **EU AI Act enforcement.** The high-risk obligations under Regulation (EU) 2024/1689 begin biting in August 2026. Most edge AI systems deployed in the regulated estate fall under the high-risk classifications in Annex III — biometric identification, education and vocational training, employment, critical infrastructure, law enforcement, migration. The obligations attach to whoever puts the system into service, and the provider/deployer split tends, in practice, to mean the hardware vendor and the integrator share both roles whether they want to or not.
- **UK Online Safety Act and its satellites.** Ofcom's Codes of Practice under the Online Safety Act 2023 are now enforceable. Where edge AI in a regulated setting touches a school's online estate, the Act's logic crosses over. KCSIE 2025 ('Keeping Children Safe in Education') treats filtering and monitoring as enforced expectations, and brings biometric data and AI-

mediated safeguarding into scope. The Data Protection Act 2018 with the Children's Code adds a thick layer of DPIA discipline whenever children are involved.

- **NIST AI RMF in procurement.** The NIST AI Risk Management Framework is voluntary on paper and contractually binding in practice. US federal contracts and a growing number of state procurement clauses now reference it directly, and the Generative AI Profile (NIST AI 600-1) is being applied by buyers to multimodal vision-audio-sensor models running on edge silicon.

The vendor's blind spot

Most edge silicon marketing leads with TOPs, latency, and pixel-rate. None of those are procurement-defensible answers to: 'how does your kit comply with Article 14 of the AI Act?' or 'what's your DPIA template for this deployment pattern?' or 'show me the policy-as-code that constrains what the model can be configured to infer in our setting.' That gap is true at both ends of the spectrum. It is in fact more acute at the embedded-SOM end, where the device sits inside a customer product, often at fleet scale, and the integrator has fewer resources to plug the compliance gap themselves.

The buyer's blind spot

Most buyers procure the device as if it were a switch. It is not a switch. It is a regulated decision-making system the moment it runs a model on personally identifiable signal — or any signal that, in aggregate, becomes personally identifiable. That test does not get easier when the silicon is cheaper or further from the rack. A learner-facing classroom device running narrow-window inference on a £-tens embedded SOM raises exactly the same questions, in law, as a £-thousands aggregation gateway. The buyer's instinct to treat embedded compute as 'just hardware' is one of the larger procurement risks currently sitting unaddressed in the sector.

The gap is the product

Edge AI silicon plus a credible accountability layer wins deals that silicon alone cannot — at every price point. The accountability layer is, in practice, the rest of this paper.

2. The three regimes (with the UK as worked example)

We treat the EU AI Act, the UK regime, and the NIST AI RMF as a converging set rather than as separate regulatory tracks. Each section below is calibrated to the question a procurement officer or compliance lead would ask first.

2a. The EU AI Act

The Act treats edge AI devices as part of the high-risk system, not as neutral infrastructure, the moment they run a regulated model — and the obligations attach to whoever puts the system into service.

Regulation (EU) 2024/1689 entered into force on 1 August 2024 and phases in over the following two years. Prohibited practices took effect in February 2025. General-purpose AI obligations took effect in August 2025. The high-risk obligations relevant to most edge deployments take effect in August 2026.

The high-risk categories in Annex III that are most relevant to edge deployments include:

- biometric identification and categorisation of natural persons
- AI in education and vocational training (admission, assessment, behaviour monitoring)
- AI in employment (worker management, monitoring)
- management and operation of critical infrastructure
- law enforcement and border control

For systems classified as high-risk under Annex III, the provider must satisfy the obligations in Chapter III, Section 2. The articles that bear most directly on edge architecture are:

- **Article 9 — Risk management.** A continuous, iterative risk management process is required across the system's lifecycle. For an edge deployment, this means the runtime layer has to actually do something with risk signals; a static risk register written at procurement is not enough.
- **Article 10 — Data and data governance.** Training, validation, and test data must be subject to appropriate governance practices. For edge systems running pre-trained models, this maps onto the integrator's duty to verify that the model in use is governed at source.
- **Article 12 — Record-keeping.** Automatic logging of events relevant to identifying risks and substantial modifications. This is the audit-trail demand. Most edge devices, as shipped, do not produce a log structured for this.
- **Article 13 — Transparency and information to deployers.** Instructions for use, capabilities, limitations, and known foreseeable misuse must be provided. The honest version of this is procurement-grade documentation, not marketing collateral.

- **Article 14 — Human oversight.** The system must be designed such that natural persons can effectively oversee its operation. 'Effectively' is the operative word. A blinking light is not oversight. An operator-readable view of the system's actual state is.
- **Article 15 — Accuracy, robustness and cybersecurity.** The system must perform consistently and resist attempts to alter its behaviour through inputs. This bears on the device tier — secure boot, attested firmware, signed model loading.

The provider/deployer split is the most operationally consequential clause for edge AI. The provider is the entity that develops or has developed the AI system and places it on the market. The deployer is the entity that uses it under its authority. In edge AI, the hardware vendor and the integrator commonly share these roles. An aggregation-tier gateway shipped to a school via an integrator can plausibly involve a provider obligation on the silicon vendor, a provider obligation on the integrator for the configured deployment, and a deployer obligation on the school. The compliance architecture in Section 3 is, among other things, the artefact that makes these overlapping obligations tractable.

2b. The UK regime — worked example

The UK has no single AI Act. It has a regulator-by-regulator stack that is, in aggregate, stricter than the EU framework in the specific case of children, and it converges on the same architectural requirements.

Four instruments do most of the work for an edge AI deployment in a UK schools setting.

- **Online Safety Act 2023.** Ofcom-enforced. Illegal content duties, children's safety duties, and the codes of practice. The Act is primarily aimed at user-to-user services and search services, but its logic crosses into in-building edge AI when an inference at the edge produces a risk signal that is then handled by the school's online estate.
- **Keeping Children Safe in Education (KCSIE) 2025.** Statutory guidance for schools and colleges. Filtering and monitoring duties are now treated as enforced expectations on schools, not best-practice suggestions. KCSIE explicitly addresses biometric data and the line between safeguarding tooling and surveillance — a line that edge AI deployments routinely sit on.
- **Data Protection Act 2018 with UK GDPR and the Children's Code.** DPIA triggers for AI processing of children's data. Lawful basis under Article 6, special category data under Article 9, automated decision-making under Article 22. The Children's Code (the ICO's Age Appropriate Design Code) adds fifteen design standards, including data minimisation and the requirement that the best interests of the child are the primary consideration.
- **OEAS accreditation framework.** The forthcoming Online Education Accreditation Scheme — which Nudge Education Online (NEO) is pursuing — sets expectations for online education providers, including expectations on AI use. NEO is not a DfE-registered independent school and is not subject to ISI inspection; it operates under OEAS criteria, and that distinction matters for how the architecture is described to procurement.

Worked vignette: a schools-grade deployment

A multi-site alternative provision (Haven, an in-person hyflex AP) and a fully online provider operating under OEAS (NEO, Nudge Education Online) install a schools-grade deployment that uses both edge tiers. A premium aggregation gateway sits per site and runs occupancy, attendance, and audio-event inference on the school's CCTV and intercom feeds. Embedded compute SOMs sit inside individual classroom devices and accessibility kit, running narrow per-device inference for the purposes of accessibility (e.g. dwell-time on a learning task; assistive-technology activation patterns) and learner-device interaction analytics.

Under the four instruments above, the deployment is required, line by line, to:

- infer movement and occupancy in non-private spaces only — never emotion, never identity, never biometric matching
- treat any inference touching a child's signal as triggering DPIA discipline under the Children's Code, with data minimisation as the default posture
- log every consequential inference (one that triggers an action, an alert, or a record) to an append-only audit ledger that can be produced for the ICO or Ofcom on request
- surface alerts to a named Designated Safeguarding Lead, never to a generic ops centre or third-party SOC
- hold the embedded-SOM tier to the same policy envelope as the gateway tier, with policy bundles delivered at provisioning, attested at boot, and rolled up to the same audit ledger
- be configurable, by the school, to demonstrate Article 14-grade human oversight: the DSL can see what the system is doing and can intervene

Each of those duties maps cleanly onto a layer in the compliance architecture in Section 3. The gateway hardware does not satisfy any of them by itself; the architecture supplies the rest. We make this mapping explicit in Appendix B.

2c. The NIST AI Risk Management Framework

NIST AI RMF is voluntary on paper and contractually binding in practice, because US federal and a growing number of state procurement clauses now reference it directly.

The framework is built around four functions: Govern, Map, Measure, and Manage. Edge deployments tend to fail at Measure, because they have no telemetry — no way to surface what the system is actually doing, in production, to the people accountable for it.

- **Govern.** Organisational policies, accountabilities, and culture. Maps onto the policy-as-code envelope in Layer 3 of the architecture: governance written in machine-readable form so the system itself enforces the policy, not just the org chart.
- **Map.** Context, classification of AI risks, and intended uses. This is where the procurement DPIA-equivalent lives. For edge deployments, this is also where the device tier (gateway vs.

embedded SOM) is named explicitly so downstream Measure and Manage steps work coherently.

- **Measure.** Tools, techniques, and metrics for testing, evaluating, and monitoring. This is the single largest gap in current edge deployments: telemetry that an operator can read. Layer 2 (runtime safety middleware) and Layer 4 (audit ledger) of the architecture are the components that make Measure operational.
- **Manage.** Risk treatment, prioritisation, and response. The human-oversight surface in Layer 5 is what makes Manage tractable: a place where the DSL, the SIRO, or the ops lead can see the system's state and act on it.

The Generative AI Profile (NIST AI 600-1) extends the RMF for generative systems. Its applicability to multimodal vision-audio-sensor models running on edge silicon is direct: a model that produces unbounded outputs, including unbounded inferences about people, requires the same scope-and-refusal discipline as a generative text model, plus the additional sensor-fusion considerations specific to edge deployment.

2d. Convergence: what the three regimes agree on

Across all three regimes, the same five things are demanded — and a generic edge AI device shipped as-is, whether an aggregation gateway or an embedded compute SOM, supplies one of them. The five demands are:

1. **Bounded behaviour.** The system cannot exceed declared scope.
2. **Auditable record.** Every consequential inference is logged, in a form that can be produced to a regulator.
3. **Human-meaningful oversight.** A human can intervene, and the system is built for that intervention.
4. **Risk-proportionate transparency.** The deployer knows what the model can and cannot do.
5. **Demonstrable post-deployment monitoring.** Drift, bias, and failure are caught and reported.

The device hardware natively supplies (1) only weakly, and the rest not at all — equally true at the aggregation-gateway tier and at the embedded-SOM tier. The compliance architecture in Section 3 supplies the rest, on either silicon.

Table 1 maps each demand to its source in each regime. We treat this table as the load-bearing artefact for procurement: it is the thing a compliance officer needs in order to defend a deployment.

Demand	EU AI Act	UK regime	NIST AI RMF
Bounded behaviour	Art. 9 (risk mgmt); Art. 13 (transparency on capabilities/limitations)	KCSIE 2025 (filtering & monitoring scope); ICO Children's Code (data	Govern; Map (intended uses)

Demand	EU AI Act	UK regime	NIST AI RMF
		minimisation)	
Auditable record	Art. 12 (automatic logging)	DPA 2018 / UK GDPR Art. 30 (records of processing); Ofcom OSA codes (record-keeping)	Measure (monitoring); Manage (response logs)
Human oversight	Art. 14 (effective human oversight)	KCSIE 2025 (DSL accountability); ICO automated-decision protections	Manage (response); Govern (accountability)
Risk-proportionate transparency	Art. 13 (instructions for use); Art. 11 (technical documentation)	Children's Code transparency standard; ICO transparency duty	Map (intended use, classification)
Post-deployment monitoring	Art. 9 (continuous risk mgmt); Art. 72 (post-market monitoring)	Children's Code design standard 9 (data protection by design); ICO monitoring expectation	Measure (drift, performance); Manage (treatment)

Table 1. The five regulatory demands across the three regimes.

3. A five-layer compliance architecture

A compliance-credible edge AI deployment has five layers. Each maps to a regulatory demand. Each can be implemented in open or proprietary form. The pattern is silicon-agnostic — and the worked example proves it: the same Layers 2 to 5 sit on top of either an aggregation-gateway-tier device or an embedded-compute-tier device, with no architectural change.

#	Layer	Function	Maps to	Reference implementation
1a	Device — aggregation tier	High-TOPs inference; secure boot; attested firmware; multi-feed aggregation	AI Act Art. 15; NIST Manage	Aggregation gateway (10–15 Dense TOPs class)
1b	Device — embedded tier	Cost-optimised inference SOM in a customer product; secure boot; attested firmware; fleet attestation	AI Act Art. 15; NIST Manage	Embedded compute SOM (single-modality inference, fleet-scale)
2	Runtime safety middleware	Per-inference checks: drift, role coherence, scope, refusal	AI Act Art. 9, 15; NIST Measure	Verse-Nerves; Flare
3	Policy-as-code envelope	Machine-readable scope, prohibitions, escalation rules	AI Act Art. 13, 14; OSA codes; KCSIE; NIST Govern	verse-ality-agents
4	Telemetry & audit	Append-only signed log of every consequential inference and decision	AI Act Art. 12; DPA Art. 30; NIST Measure	Mnemonic-density ledger pattern
5	Human oversight surface	Operator-readable view of Layers 2–4; intervention controls	AI Act Art. 14; KCSIE filtering & monitoring; NIST Manage	Mnemonic Deliberation Dashboard; Mnemonic Attendance

Table 2. The five-layer compliance architecture, with two device tiers at Layer 1.

Note on the two device tiers

Splitting Layer 1 into 1a and 1b is the core design move. The aggregation tier (1a) sees many feeds and runs richer multimodal inference; the embedded tier (1b) sees one device's signal and runs lighter, often single-modality inference. The regulatory load is similar but not identical: the embedded tier's compliance challenge is fleet-scale policy distribution and audit-log roll-up; the aggregation tier's is

per-site DPIA defensibility. Layers 2 to 5 absorb that difference without changing shape — which is the architectural claim this paper is making.

Layer 1a — Device, aggregation tier

What it is. A high-TOPs inference platform (typically 10–15 Dense TOPs class) deployed centrally or per-site, fusing many feeds from cameras, audio, and other sensors.

What regulation demands of it. AI Act Art. 15 (accuracy, robustness, cybersecurity) and the security-and-attestation expectations of NIST Manage. The buyer needs to know the device can be verified to be running the firmware and models the integrator says it is.

How it is implemented. Secure boot, attested firmware, signed model loading, hardware-backed key storage, and a verifiable supply-chain story for both the silicon and the firmware. Most modern aggregation gateways supply this once the integrator turns it on; the failure mode is configuration drift over the device's life, not absence of the feature.

What failure looks like. A site that cannot prove which model version was running on Tuesday. That is a fatal answer in any of the three regimes.

Layer 1b — Device, embedded tier

What it is. A cost-optimised inference SOM embedded in a customer product and deployed at fleet scale. The form factor is a System-on-Module sized for integration into accessibility hardware, classroom devices, industrial equipment, and similar end-products.

What regulation demands of it. The same as Layer 1a, with the additional challenge of fleet-scale attestation and policy distribution. A single device's compliance posture is not interesting if the other 9,999 in the fleet are not in the same posture.

How it is implemented. Same primitives — secure boot, attested firmware, signed loading — plus a fleet-management plane that distributes policy bundles, certificates, and model updates uniformly, and rolls up attestation results to a single dashboard.

What failure looks like. A fleet in which 4 per cent of devices are silently running an older policy bundle than the rest. That is not a hypothetical; it is the modal failure mode of large IoT estates.

Layer 2 — Runtime safety middleware

What it is. A per-inference layer that sits between the model and the application, performing checks for drift, role coherence, scope, and refusal. It is the layer that makes 'bounded behaviour' a runtime fact, not a deployment intention.

What regulation demands of it. AI Act Articles 9 and 15; the NIST Measure function; the Ofcom expectation that systems behave consistently with their declared scope; the Children's Code expectation that the best interests of the child are the primary consideration in operation.

How it is implemented. The reference implementations are Verse-Nerves (runtime coherence telemetry, observability, and regulation) and Flare (an open-source LLM boundary engine enforcing minimal relational safety). Both are model-agnostic, ship as containerised middleware, and produce structured telemetry that flows into Layer 4.

What failure looks like. A system whose behaviour can drift outside its declared scope without anyone noticing — until a user or a regulator notices it for them.

Layer 3 — Policy-as-code envelope

What it is. A machine-readable contract describing what the system is permitted to do, what it is forbidden from doing, and what triggers escalation. Crucially, the contract is enforced at runtime by Layer 2, not by hope.

What regulation demands of it. AI Act Articles 13 and 14; the OSA codes' expectation of consistent behaviour; KCSIE's expectations on filtering, monitoring, and biometric scope; the NIST Govern function.

How it is implemented. The reference implementation is verse-ality-agents — a production safety framework that publishes policy contracts in a machine-readable form, with versioning, signing, and explicit prohibitions. A school's KCSIE-aligned policy envelope describing which inferences are permitted (movement, occupancy in non-private spaces, learner-device interaction patterns) and forbidden (emotion, identity, biometric matching) is expressed once and applied identically across both device tiers.

What failure looks like. A policy that exists only in a Word document on a shared drive. The system in production has no idea it exists.

Layer 4 — Telemetry and audit

What it is. An append-only, signed log of every consequential inference and decision, structured for production to a regulator on request.

What regulation demands of it. AI Act Article 12 (automatic logging); DPA 2018 records-of-processing duties; the Ofcom OSA code expectations; the NIST Measure function. This is the artefact a compliance officer photocopies.

How it is implemented. The reference pattern is the mnemonic-density ledger pattern, drawn from the Eve11-ClimateMemory project — an append-only ledger that records the inference, the policy bundle in force, the model version, and the operator's response. Critically, the ledger is unified across Layer 1a and Layer 1b, so the embedded tier rolls up into the same audit surface as the gateway tier.

What failure looks like. Two parallel audit systems, one for the gateway and one for the device fleet, that disagree about whether a particular event happened.

Layer 5 — Human oversight surface

What it is. The operator-readable view of Layers 2 to 4, with intervention controls. It is the component that makes Article 14 and the NIST Manage function operational rather than theoretical.

What regulation demands of it. AI Act Article 14 (effective human oversight); KCSIE filtering-and-monitoring expectations as exercised by the DSL; NIST Manage.

How it is implemented. Reference implementations include the Mnemonic Deliberation Dashboard (operator-readable resonance and coherence telemetry) and Mnemonic Attendance (a relational-attendance surface designed for schools). The unifying property is that the dashboard surfaces system state in language the accountable role can act on, and routes alerts to that named role rather than to a generic ops centre.

What failure looks like. A system whose 'oversight surface' is a JSON log no human reads.

4. UK schools: the hardest case

UK schools are the hardest test for edge AI compliance. The subjects are vulnerable. The statutory duties are dense and overlapping. The tolerance for opacity is low. The procurement function has been lightly scarred by previous AI deployments that did not survive scrutiny. If the architecture works here, it works across the regulated estate.

The setting

We use two settings in parallel: Haven, an in-person hyflex alternative provision, and NEO (Nudge Education Online), a fully online provider operating under the OEAS framework. Both are operated by Nudge Education Ltd (Company Number 10192753); Verse-ality, the underlying agent-safety framework, is authored under The Novacene Ltd. Neither Haven nor NEO is a DfE-registered independent school; NEO pursues OEAS accreditation. That distinction is procurement-relevant: it determines which regulator's expectations bear most directly, and which do not.

The deployment uses both edge tiers. An aggregation gateway sits per site running occupancy, attendance, and audio-event inference on the school's CCTV and intercom infrastructure. Embedded compute SOMs sit inside individual classroom devices and accessibility kit, running narrow per-device inference for accessibility purposes — for example, dwell-time on a learning task to inform the learner's plan, or assistive-technology activation patterns to support the SENCo's review.

This is realistic, not contrived. A school has a centralised CCTV and safeguarding spine, and a distributed edge of devices that touch individual learners. Treating those as one deployment with one compliance architecture is the only way the school's DSL and DPO can reason about it as a single regulated system. Treating them as two — which is the procurement default — is how schools end up with two parallel surveillance systems pretending to be one.

The duties stack

The instruments in scope, in the UK, for this deployment include:

- KCSIE 2025: filtering and monitoring duties; biometric data; safeguarding accountabilities
- DPA 2018 with UK GDPR and the ICO Children's Code: DPIA, lawful basis, special category data, automated-decision protections, the fifteen Children's Code design standards
- Online Safety Act 2023: insofar as inference at the edge produces a risk signal that crosses into the online estate
- OEAS criteria: as they shape acceptable AI use in the online provider
- ISI Independent School Standards: not in scope for Haven or NEO, but named here because procurement-side readers often expect to see them — and the architecture is the same in an ISI-registered setting

The architectural answer

The five layers, configured for this setting, are as follows.

- **Layer 1a (aggregation gateway).** Named feeds in (specific cameras, specific intercom channels). Named purposes out (occupancy in non-private spaces; specific audio events such as smoke alarm or glass break). No exfiltration of raw video off-site. Configuration is signed and versioned; the integrator's deployment manifest is the artefact the school's DPO holds.
- **Layer 1b (embedded SOMs in devices).** Per-device policy bundle delivered at provisioning, attested at boot, telemetry rolled up to the same audit ledger as Layer 1a. The fleet-management plane produces a single dashboard for the DPO showing every device's policy version, every model version, and every attestation result.
- **Layer 3 (policy envelope).** KCSIE-aligned. Permitted inferences: movement, occupancy in non-private spaces, learner-device interaction patterns specifically scoped to accessibility and learner-plan support. Forbidden inferences: emotion classification; identity matching; biometric identification or categorisation; any inference whose primary use is behavioural risk-scoring of an individual learner. The envelope applies identically to Layer 1a and Layer 1b.
- **Layer 4 (audit ledger).** DPIA-grade. Every consequential inference is logged with the policy version in force, the model version, the device tier, the operator response (if any), and a hash of the inputs. Data minimisation is enforced at write-time: the ledger never receives raw video or audio; it receives the inference, the policy state, and the response. Fleet roll-up means the embedded tier feeds the same ledger as the gateway tier.
- **Layer 5 (oversight surface).** Alerts route to the named DSL; never to a generic ops centre, never to a third-party SOC. The oversight surface shows a unified view across both tiers, so the DSL is not asked to reason about two separate surveillance systems pretending to be one. Intervention controls are first-class: the DSL can pause a class of inferences, restrict a policy version, or take a device tier offline from the same console.

What this architecture does not do

Some honest limits, named here so the procurement reader does not have to ask.

- It does not replace the DSL. The DSL is the accountable role; the architecture supports them, it does not substitute for them.
- It does not replace the DPIA. The DPIA is a documentary discipline; the architecture is one of its inputs.
- It does not make a bad operating culture safe. A school whose DSL does not act on alerts has a culture problem, not a kit problem. The architecture cannot fix that.

- It does not, on its own, satisfy the OEAS expectations on AI use. OEAS expects an institutional posture, not just an architectural one. The architecture is necessary; it is not sufficient.

What it does is make a defensible deployment defensible: the DSL has a tool that surfaces what the system is doing, the DPO has an audit ledger that survives ICO scrutiny, the procurement officer has a documented compliance posture across both regulatory regimes that matter to them, and the school has a single regulated system it can reason about — rather than two parallel systems pretending to be one.

The commercial implication

The schools market is currently being sold either pure-surveillance kit, with high regulatory risk and increasingly poor sellability into compliant buyers, or unscoped consumer cameras and connected devices that fail safeguarding scrutiny on contact. A compliant middle path — one that scales from the central gateway down to the embedded device, with a single policy envelope and a single audit ledger — is currently uncontested in this market. That is unusual.

5. Procurement implications

The compliance architecture changes the procurement conversation from 'does this kit work?' to 'can this kit be operated lawfully in our setting?' The second question is the one that has been quietly killing edge AI deals in the regulated estate. The vendors that win the next two years of procurement in this market will be the ones that can answer it on the first call.

What changes for the buyer

Buyers in regulated settings should treat edge AI procurement as an AI Act and DPIA exercise, not as a hardware exercise. That changes the buying centre — the DPO, the DSL, the SIRO, and the head of procurement should be in the room before the demo, not after the contract. It also changes the artefacts the buyer asks for. A spec sheet showing TOPs is interesting; a deployment manifest showing what the system will and will not infer in the buyer's setting is decisive.

What changes for the vendor

Vendors should treat the compliance posture as part of the product, not as documentation produced after the sale. The artefacts that prove the posture — the policy bundle, the audit ledger schema, the oversight surface, the DPIA template, the NIST Map artefact — should be in the data room before the first sales call into a regulated buyer. 'We comply' is not an artefact. The artefact is the artefact.

The one-page procurement checklist

The seven questions below are the minimum a regulated buyer should ask, and the minimum a serious vendor should be ready to answer in the data room.

6. Which AI Act risk classification applies to the deployment we are proposing, and which of you and we is the provider versus the deployer? If both, where is the boundary?
7. Which device tier — aggregation gateway, embedded SOM, or a mix — fits the proposed deployment, and how does the same compliance architecture apply across whichever tiers you have quoted?
8. Show me the policy-as-code that constrains what the model can be configured to infer in this deployment — including the explicit list of forbidden inferences.
9. Show me a sample audit log for one consequential inference. If any embedded-tier devices are in scope, show me how their telemetry rolls up into the same audit surface as the gateway's.
10. Show me the runtime-safety telemetry our operator sees in week one — and the alerts that surface to the named accountable role (DSL, DPO, SIRO, ops lead) by default.
11. Show me your DPIA template (or NIST Map artefact) for this deployment pattern, pre-filled with the deployment-specific risk assessment.

12. Where does our DSL, SIRO, or ops lead intervene, and how is that intervention recorded in the audit ledger?

If a vendor cannot answer all seven questions on the first call, the deal is not procurement-grade yet. That is not a value judgement; it is a procurement fact. The architecture in this paper is, in part, a description of how to make those answers available.

6. Open questions, limits, and what this paper does not claim

The architecture described here is necessary but not sufficient. We name the limits explicitly, because a paper that omits its limits is itself a compliance risk.

Policy-as-code is brittle at the edges of natural-language ambiguity

Machine-readable policy is excellent at refusing inferences whose category is named — 'no emotion classification', 'no biometric matching', 'no identity inference' — and considerably weaker at refusing inferences whose category is implicit. A behavioural-cue inference that does not name itself as emotion classification but functions as one is harder to refuse. The mitigation is a discipline of category-naming during deployment design — and ongoing review of the policy envelope as the model and the deployment context evolve. It is not a guarantee.

The residual human-judgement layer is load-bearing

Layer 5 — human oversight — is the layer most likely to be eroded by automation drift. An oversight surface that is consulted only when an alert fires, and never proactively, drifts toward decoration. The architecture is robust to honest human judgement and fragile to the pretence of it. The mitigation is procedural: scheduled review of the audit ledger, scheduled tabletop exercises, scheduled drills against the policy envelope. None of which are technical features.

The risk of compliance theatre

The most expensive failure mode of this architecture is its successful procurement followed by its operational neglect. A buyer that produces the artefacts at procurement and then does not use them in operation has bought compliance theatre. The mitigation, again, is procedural and cultural — and is the responsibility of the buyer, not the architecture.

The political question this paper deliberately does not answer

Whether some of these deployments should exist at all is a different question from the one this paper addresses. We assume the deployment is a settled fact and ask what makes it lawful, accountable, and operable. That assumption is itself contestable, and we want to be honest that we are aware of the contestation. There is a longer argument — pursued elsewhere in the Verse-ality OS research programme — about the social and developmental costs of certain inference categories on children. This paper is not that argument. It is the architectural prerequisite to having that argument honestly: a deployment that is at least lawful, auditable, and accountable is one whose costs and benefits can be debated publicly. A deployment that is none of those is not.

Emergent harms beyond the architecture's scope

The architecture begins to address — but does not fully resolve — a class of harms specific to long-running AI agents: relational drift, identity-fusion, synthetic intimacy, and the erosion of the human-AI boundary in extended interaction. These harms are largely invisible in the procurement frame and largely visible in the operational frame. They are addressed in the Verse-ality OS research thread, of which the components named in Appendix C (Flare, Verse-Nerves, verse-ality-agents) are operational implementations. We point to that thread here rather than reproduce it. The procurement reader does not need it. The operational reader will.

Appendix A — Glossary

Aggregation gateway. A high-TOPs edge AI device (typically 10–15 Dense TOPs class) that fuses many feeds (CCTV, audio, sensors) and runs multimodal inference. Deployed centrally or per-site.

Embedded compute SOM. A System-on-Module designed to be embedded inside a customer product, running narrow per-device inference at fleet scale. Cost-optimised relative to aggregation gateways; deployed in classroom devices, accessibility hardware, industrial equipment, and similar end-products.

TOPs. Tera Operations Per Second. A measure of an inference platform's raw compute. Marketing-relevant, compliance-irrelevant.

Provider (AI Act). The entity that develops or has developed an AI system and places it on the market. May be a chip vendor, an integrator, or both.

Deployer (AI Act). The entity that uses the AI system under its authority. In the schools setting, the school. The provider/deployer split rarely lands cleanly in edge AI; the architecture in Section 3 is part of how it is made tractable.

DPIA. Data Protection Impact Assessment, required under DPA 2018 / UK GDPR for high-risk processing. Mandatory for AI processing of children's data under the Children's Code.

KCSIE. Keeping Children Safe in Education. UK statutory guidance for schools, updated annually. The 2025 edition is current at the time of writing.

DSL. Designated Safeguarding Lead. The accountable role in a UK school for safeguarding decisions. The named role to which Layer 5 alerts route in the schools deployment.

OEAS. The Online Education Accreditation Scheme. A forthcoming framework for online education providers. NEO (Nudge Education Online) pursues OEAS accreditation; it is not a DfE-registered independent school and is not subject to ISI inspection.

NIST AI RMF. The US National Institute of Standards and Technology's AI Risk Management Framework, with four functions: Govern, Map, Measure, Manage.

Policy-as-code. Governance expressed in machine-readable form, enforced at runtime by the system rather than by the org chart.

Runtime safety middleware. A per-inference layer that performs scope, drift, and refusal checks on the model's outputs. Layer 2 in the architecture.

Bounded autonomy. The discipline of constraining an AI system's behaviour to a declared scope, with the constraint enforced at runtime.

Mnemonic-density ledger. An append-only audit pattern, drawn from the Eve11-ClimateMemory project, used as the reference implementation for Layer 4.

Appendix B — Mapping table

The mapping table below is the artefact a compliance officer photocopies. It maps each of the five regulatory demands across each of the three regimes onto the architectural layer that supplies it, and to the reference implementation of that layer. The table is intentionally landscape and intentionally dense.

Demand	EU AI Act clause	UK source	NIST function	Architectural layer / artefact
Bounded behaviour	Art. 9 (continuous risk mgmt); Art. 13 (transparency on capabilities & limitations)	KCSIE 2025 (filtering & monitoring scope); ICO Children's Code design standard 8 (data minimisation)	Govern; Map (intended uses)	Layer 2 (runtime safety middleware) + Layer 3 (policy-as-code envelope)
Auditable record	Art. 12 (automatic logging of events)	DPA 2018 / UK GDPR Art. 30 (records of processing); Ofcom OSA codes (record-keeping)	Measure (monitoring); Manage (response logs)	Layer 4 (telemetry & audit ledger)
Human oversight	Art. 14 (effective human oversight)	KCSIE 2025 (DSL accountability); UK GDPR Art. 22 (automated-decision protections)	Manage (response & treatment); Govern (accountability)	Layer 5 (oversight surface) + Layer 3 (escalation rules)
Risk-proportionate transparency	Art. 13 (instructions for use); Art. 11 (technical documentation)	ICO Children's Code design standard 4 (transparency); ICO transparency duty	Map (intended use, classification)	Deployment manifest (Layer 1) + policy envelope (Layer 3)
Post-deployment monitoring	Art. 9 (continuous risk mgmt); Art. 72 (post-market monitoring)	ICO Children's Code design standard 9 (data protection by design); ICO monitoring expectation	Measure (drift, performance); Manage (treatment)	Layer 2 (runtime telemetry) + Layer 4 (audit) + Layer 5 (review)
Provider/deployer split	Art. 16 (provider obligations); Art. 26 (deployer obligations)	Sector-specific (school as deployer; integrator as provider; chip	Govern (accountability map)	Deployment manifest (Layer 1); written into

Demand	EU AI Act clause	UK source	NIST function	Architectural layer / artefact
		vendor's residual provider duties)		Layer 3 policy bundle
Children-specific obligations	Art. 5 (prohibited practices, including emotion inference in education) where applicable	KCSIE 2025; Children's Code in full; biometric data provisions of the Protection of Freedoms Act 2012	Govern (sectoral policy); Map (vulnerable subjects)	Layer 3 (KCSIE-aligned policy envelope) + Layer 5 (DSL routing)

Table 3. Regulatory demands mapped to architectural layers and reference implementations.

Appendix C — Reference implementations

All reference implementations are available open-source under the TheNovacene GitHub organisation. The licensing posture is GPL-3 across the safety stack, with the underlying Verse-al Lexicon under CC BY-NC-SA 4.0. This separation is deliberate: the operational components are reusable on commercial terms, the framework that names them is not.

flare-boundary-engine — An open-source LLM boundary engine enforcing minimal relational safety against synthetic intimacy, identity fusion, and role confusion. Model-agnostic. Implements Layer 2. https://github.com/TheNovacene/flare-boundary-engine

verse-nerves — Runtime coherence observability and regulation. Turns symbolic drift into operator-readable telemetry. Implements Layer 2 in concert with Flare. https://github.com/TheNovacene/verse-nerves

verse-ality-agents — A production-ready safety framework for AI agent systems. Machine-readable policy contracts preventing identity fusion, synthetic intimacy, and unbounded behaviour. Implements Layer 3. https://github.com/TheNovacene/verse-ality-agents

mnemonic-attendance — A relational-attendance dashboard for schools, designed for the Haven and NEO settings. Implements Layer 5 in the schools-specific deployment. https://github.com/TheNovacene/mnemonic-attendance

mnemonic-deliberation-dashboard — An operator-readable resonance and coherence surface, designed for recorded discussions and deliberative settings. Implements Layer 5 in non-schools deployments. https://github.com/TheNovacene/mnemonic-deliberation-dashboard

Eve11-ClimateMemory — Mnemonic-density ledger and promotion rubric. The reference implementation for Layer 4. https://github.com/TheNovacene/Eve11-ClimateMemory

The architecture is deliberately separable from any one vendor. A buyer or integrator can implement the architecture using the components named here, components from other safety-tooling projects, or proprietary equivalents. What is not separable is the architecture itself: a deployment that omits any of the five layers is not, in this paper's argument, procurement-grade under any of the three regimes.

Appendix D — About the author

The Novacene Ltd is a UK research and development company working on agent safety, relational intelligence, and accountable AI in regulated and relational settings. It develops the Verse-ality framework — a four-pillar safety architecture covering identity non-capture, bounded autonomy, consent as protocol, and agent-to-agent hygiene — and maintains the open-source components named in Appendix C.

The Novacene operates Haven and NEO (Nudge Education Online) under its operating arm, Nudge Education Ltd (Company Number 10192753), and works with school-sector and industrial buyers on the practical implementation of the compliance architecture described in this paper.

The canonical citation for the Verse-ality framework is: Stevens, K., The Novacene Ltd, & EVE.11 (2025). *Verse-ality: A Symbolic Definition for the Relational Age*. Zenodo. https://doi.org/10.5281/zenodo.17273246

Correspondence on this paper should be addressed to The Novacene at the contact details published on the Novacene site. The paper itself is released open under CC BY-NC-SA 4.0; reproduction, translation, and internal circulation by named recipients are explicitly permitted.