Fitness to Serve

What qualifies an AI system to occupy a relational role with a child

Companion note · August 2026

The third question

A school has a registrar, an address, and staff who have been checked. Enrollment asks who may enter; siting asks what a child crosses to get there; fitness asks who may work inside. This note concerns the third. Nothing in it depends on a namespace: the duties apply to a system deployed anywhere, and are expressible as admission criteria where an admission layer exists.

Fitness, not age

At a liquor store the customer proves age — one bit, tested once. At a school the adult proves fitness: checked, credentialed, trained, renewed, revocable, while the child proves nothing.

Not every system near a child inherits a school-employee regime. But as a system moves from publication toward sustained, personalised, memory-bearing interaction with an identified child, the governance question increasingly resembles fitness rather than age-gating alone.

The missing third state

Systems answer or decline. Institutions answer, decline, or refer. A refusal ends the deployer's exposure; a referral serves the child, and a child who discloses distress to a system that only declines has been left alone at the moment of disclosure.

As admission criteria the floor resolves on three independent axes — response, escalation, accountability — each declarable and checkable on audit. A bound that cannot be tested is not a bound.

Two constraints that matter most

Escalation must never be fixed to the guardian: where the parent is the source of harm, notification converts a disclosure of abuse into a disclosure to the abuser. And the system is not a statutory reporter but an internal conduit — it routes to a designated human who forms or declines to form suspicion.

A referral also has to mean something: a named reachable destination, minimum necessary context, and a record that the transfer occurred. Signposting is not referral.

Enrollment enables the fuller duty set

Escalation presupposes a reachable adult; referral presupposes a destination. An anonymous service can decline and signpost, but cannot discharge duties depending on knowing who the child is. Institutional deployments already hold that relation.

Why this could be rejected

Liability and capacity are the real obstacles. A flag may engage duty of care the moment it is raised, while a miss supplies a negligence theory; and safeguarding staff are already stretched.

Referral capacity is part of deployment capacity — but need not sit inside the deploying organisation, or the standard would hand broad roles to well-staffed schools and stripped tools to under-resourced ones. Shared district or regional capacity can carry triage; the statutory decision stays local.

What is not known

Escalation volumes, false-positive rates, re-attestation intervals and workable shared-capacity ratios are unmeasured. Whether conventional safety evaluation measures referral at all is a hypothesis, not a finding.

The full note sets out the three-axis model and four questions a trial would need to answer.

To request the full paper, write to shknudson@ifcsis.org.