

Is your AI agent ready for production?

An independent controls review for boards approving agentic AI

The question you are being asked. A team wants to put an AI agent into production. It will read documents, call systems and take actions on the firm's behalf, at a rate no human can review. The paper in front of you says a human is in the loop and the policy is in place. You are being asked to approve go-live on that basis.

Why the policy is not the answer. When something goes wrong, the regulator will not ask what the policy said. Under SMCR, the named Senior Manager will be asked what actually controlled the agent, what it could reach, and who could stop it. Consumer Duty asks the same about outcomes. An insurer will ask whether the questionnaire the firm completed matched the platform it was running. In our experience the gap between the policy and the platform is where the incidents live.

What the review does. The Gordion Agentic AI Health Check tests the running platform, not the document, against thirty controls in six areas: whether model output can execute directly; what the agent can reach and with what credentials; whether budgets, sequencing and logging are enforced in code; who owns it and reads the logs; what data it sees and can leak; and what it depends on that the firm does not control. Every control is graded Verified, Partial, Not evidenced or Contradicted, with the artefact that proves it.

Two tests are done live. The kill switch is pressed in production, under written authorisation, by the named non-engineer, and timed. Prompt injection is tested on every surface the agent reads, including the case nobody looks for: a legitimate action with a poisoned value.

What you receive. A four-page board paper: the go-live opinion, the three things to fix, the loss scenarios the platform can trigger and which are bounded, and a one-page Statement written for insurers and supervisors. Behind it, a full findings pack with a 90-day remediation plan. The assessor presents to the board.

What it is not. It is not a certification, a warranty, or an assessment of whether the agent is any good at its job. It is a point-in-time controls review of one platform, by an assessor who does not fix what is found during the review.

Time and cost. Four weeks from engagement letter to board decision. Fixed fee, quoted at scoping.

Next step. Ask your platform owner to complete the free twenty-question Readiness Check at gordionsolutions.co.uk/readiness-check. It records what the firm believes is true. The Health Check tells you which of those beliefs hold.

Antony Coppellotti, Gordion Solutions. Founder and CTO.