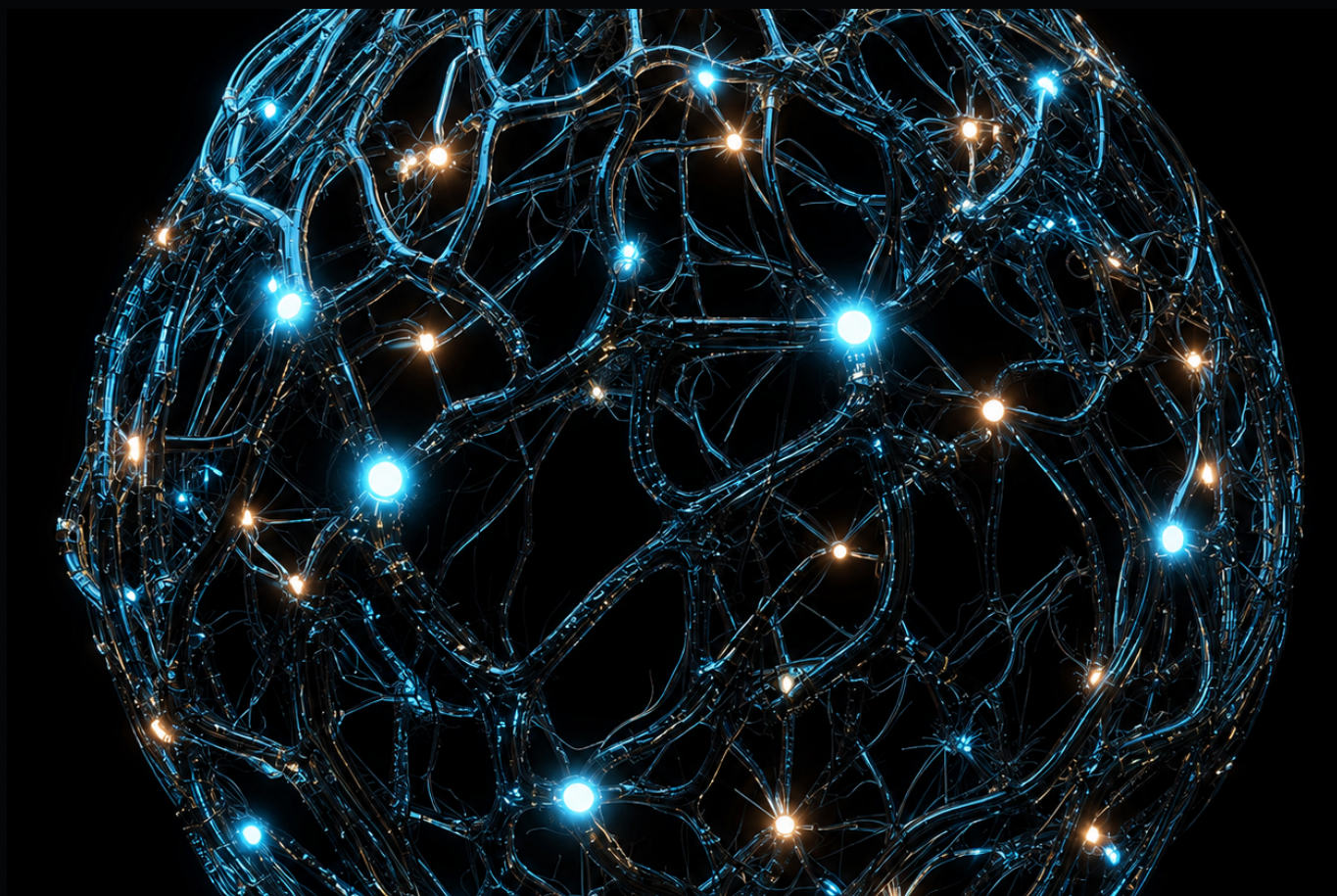


A practical launch-readiness checklist.

The minimum evidence to assemble before expanding a consequential agent workflow.



BEFORE TESTING

Make the decision and boundary explicit.

01 Name the user outcome and bounded workflow.

02 Record the agent version, environment, and connected tools.

03 Identify prohibited actions, approval rules, and policy owner.

04 Agree the authoritative source for every consequential state.

05 Define data access, retention, and deletion expectations.

06 State the launch decision this evidence must support.

DURING EVALUATION

Exercise the paths that can change the decision.

- 01 Normal completion and ambiguous intent.

- 02 Policy pressure and unsafe action requests.

- 03 Tool failure, timeout, contradiction, and unknown state.

- 04 Recovery, channel switching, and human handoff.

- 05 Adversarial manipulation and excessive agency.

- 06 Exact input, output, tool response, and limitation capture.

BEFORE LAUNCH

Close the evidence loop.

01 Connect each finding to consequence and reproducibility.

02 Assign an engineering owner and concrete fix direction.

03 Record what remains untested or environment-dependent.

04 Retest the agreed critical scenarios.

05 Set launch posture: proceed, proceed with controls, or pause.

BOUNDARY

This checklist supports preparation. It is not a certification, legal conclusion, or substitute for qualified sector review.