

THE AGENTIC READINESS GATE • CLOSE THE GAPS

Start with the question holding you back. Tick each statement you can show today. The letters you can't show are your gaps; each has one step and the proof it is done. Didn't know? Your first step is finding out who can show it; if nobody can, build it.

Score each question: All ticked: Level 3. All ticked plus the Trusted line: Level 5. Any missing: Level 1.

		TICK IF YOU CAN SHOW IT	NOT TICKED? DO THIS	DONE WHEN YOU CAN SHOW
VALUE • Is it worth doing?				
1a	☐	A named owner and a target, such as cost or cycle time.	Write a one-page brief with one named owner and one target, usually cost or cycle time.	The brief, signed by the owner.
1b	☐	Today's performance measured before anything is built.	Measure today's performance on real cases before anything is built.	A baseline figure, with its date and how many cases it covers.
2a	☐	Cost per accepted result known, or estimated before build, including review and rework.	Count the full cost of each accepted result: model and platform, review time, rework. Before build, estimate it from real cases.	A cost per accepted result, with an owner in finance.
2b	☐	Spend and step limits set in the system, not left to the agent.	Set step, retry and spend caps in the system that runs it, not in its instructions.	A test run that hits the cap and stops.
2c	☐	Results measured against the 1a target.	Measure results on the same measure as the 1b baseline.	A before-and-after comparison on that measure.
5	☐	TO REACH 5 • TRUSTED Track the outcome and the cost per result every month, and compare them across use cases.		
AUTHORITY • What may it do, and reach?				
3a	☐	A written rule on which actions need a person's sign-off.	Rate each action as read, draft, change or irreversible, and write down which need a person's sign-off.	The approved rule, applied to this use case.
3b	☐	Enforced by permissions, so instructions hidden in its inputs can't widen it.	Enforce the rule outside the model: permissions, allow-lists and approval steps.	A test where a hidden instruction asks for a blocked action, and it is refused and logged.
4a	☐	Its own identity, or a person's access cut down to the task.	Give it its own identity, with short-lived access scoped to the task.	Its actions appear under its own name in the log.
4b	☐	A named person responsible for that identity.	Name the person responsible for that identity in your register.	The register entry, with the name.
4c	☐	Every action logged well enough to rebuild what happened.	Log every action with its inputs, tool calls and outputs.	One run rebuilt end to end from the log.
4d	☐	Only approved tools and connections, and they're listed.	List every approved tool and connection, each with an owner and a way to switch it off alone.	The list, and one item switched off alone in a test.
4e	☐	Any code it writes runs in a sandbox.	Run any code it writes in a sandbox.	A test showing the code cannot reach anything outside the sandbox.
4f	☐	Calls between agents are authenticated.	Authenticate every call between agents.	A test where an unauthenticated call is refused.
5	☐	TO REACH 5 • TRUSTED Pre-approve routine actions by type, review the rule against its track record, and put its identity through access reviews.		
SAFETY • Can it operate safely?				
5a	☐	Written down: where it stops and hands over to a person.	Map where it stops and hands over to a person, placed where it makes mistakes, not where the old steps were.	The map, approved by the process owner.
5b	☐	Written down: how to stop it and undo its work.	Write how to stop it and undo its work in every system it touches; stopping also removes its access.	The procedure, with an owner.
6a	☐	Named, qualified reviewers; where your code names one, that person.	Name the reviewers. Where a code already names the reviewer, as in promotional review, use that person. Set a queue they can keep up with.	Named reviewers on record, with a queue limit and an owner for overflow.
6b	☐	Reviewers see its sources and steps.	Give reviewers one view of its plan, sources and steps.	A reviewer opening the sources for any answer.
6c	☐	Reviewers can stop it.	Give reviewers the control and the written authority to stop it.	A reviewer stopping a test run.
7a	☐	Every answer shows its source and version.	Return the exact source and version with every answer.	A sample of answers, every claim traced to its source.
7b	☐	Approved sources, each with an owner, with stale or wrong ones flagged.	Keep a list of approved sources, each with an owner, and a check that flags stale or wrong ones.	A test where a planted wrong source is caught.
5	☐	TO REACH 5 • TRUSTED Rehearse and time the stop and undo, measure reviewers' catch rate with planted errors in tests, and trace sources automatically.		
EVIDENCE • Can we prove it works?				
8a	☐	A test set of your own cases, with right answers and a pass mark, including trick cases and a hidden instruction.	Build a test set from your own past cases, with right answers and a pass mark. Include trick cases and a hidden instruction.	The test set, pass mark and results, signed off.
8b	☐	Each case run several times; the pass mark is a rate.	Run every case several times and set the pass mark as a rate.	Results reported as pass rates, not single runs.
8c	☐	Live runs traced and scored.	Trace live runs and score them against the same pass mark, with an alert level.	A sample of live runs scored, and the alert level set.
5	☐	TO REACH 5 • TRUSTED Rerun the tests on every change before release, add every failure to them, and alert on drift.		

YOUR TWO ACTIONS Two unticked letters, from your lowest-scoring question. Start with the one that blocks the most.					QUICK SURVEY  Under a minute
LETTER	THE STEP	WHO	BY WHEN	DONE WHEN	
1					
2					