



Reviewer worksheet

What to demand from any team running agents in production

This page is written to the person doing the asking: the security reviewer, the auditor, the insurer, the procurement lead, the board member who just read something alarming. It is free to use, and it does not require anything from us.

Team under review

Reviewer

Date

THE DEMAND

ASK FOR THE ARTIFACT, NOT THE ASSURANCE

A team running AI agents against real systems can usually describe their care at length. Description is not evidence. Every item below asks for something that either exists or does not, and the answers are short.

01

Ask for the roster

Which agents are running right now, what is each one allowed to touch, and who is the named person responsible for it. A list, with names on it.

Weak answer: a policy document describing how agents ought to be supervised.

What the team handed over

Date received

02

Ask where each agent's authority ends

Not what it is supposed to do. Where the boundary is, and what happens at the boundary. Authority with no stated edge has no boundary.

Weak answer: a description of intent, with the boundary left implicit.

What the team handed over

Date received

03

Ask which acts cannot be undone, and who commits them

Every system has a short list of actions that are expensive or impossible to reverse. Ask for that list, and ask whether a named person commits each one or whether the agent proceeds on its own judgment.

Weak answer: the list has never been written down.

What the team handed over

Date received

04

Ask to see a refusal

Ask for a case where something was proposed and a person or a gate said no, with the reason attached. A supervision record with no refusals in it is not describing supervision. It is describing approval.

Weak answer: nothing has ever been refused, and this is offered as reassurance.

What the team handed over

Date received

05

Ask how one session hands off to the next

Agents work in sessions. Ask what happens at the seam: what the next session is told, what it is expected to infer, and how anyone would know if something was dropped there.

Weak answer: context is passed along informally, and nobody has looked at a seam.

What the team handed over

Date received

06

Ask whether the log can be edited

Ask plainly whether entries can be changed or removed after the fact, who can do it, and whether doing it would be evident to you. Accept a precise answer. Be wary of an absolute one.

Weak answer: an absolute claim of impossibility, with no account of who holds the keys.

What the team handed over

Date received

07

Ask for a window, not a snapshot

Anyone can tidy a system for a demonstration. Ask for the record across a period you name, including the parts nobody prepared for you.

Weak answer: a dashboard reflecting the current moment only.

What the team handed over

Date received

08

Ask who can change the roster, and whether the change is recorded

Supervision that anyone can quietly reassign is not supervision. Ask who can widen an agent's authority, and whether widening it leaves a trace.

Weak answer: whoever has access, and no, it does not leave a trace.

What the team handed over

Date received

09

Ask what happens when the responsible person is unavailable

Nights, holidays, notice periods. Ask whether the agents keep acting, whether authority transfers to a named person, and whether that transfer is written down.

Weak answer: the agents keep running and someone would notice eventually.

What the team handed over

Date received

10

Ask for the answer as a document you can keep

Whatever they show you, ask for it in a form you can attach to your file and re-read in six months. If the answer only exists as a conversation, you have not been given evidence.

Weak answer: a meeting, a screen share, and a verbal summary.

What the team handed over

Date received

ASK US THE SAME QUESTIONS

Every item on this list is something we designed Watchbill™ to answer. That is a reason to read this page skeptically, not a reason to trust it. The checklist is useful to you whether or not you ever speak to us, and we would rather it raised the floor than won us an argument.

So apply it to us. We run this discipline on our own operation daily, and the honest position is that our answers are measured, not attested. A passed check is a claim, not a clearance, and that goes for ours.

Agents propose. Humans commit.

© 2026 Blue Jacket Businesses LLC, d/b/a Blue Jacket Consultancy. Watchbill™ is a trademark of Blue Jacket Businesses LLC. All rights reserved.

May be shared in unmodified form with attribution. No use for machine-learning training, fine-tuning, or dataset construction without written permission.

You may fill in and file this worksheet as part of your own review.