

AI Agent Readiness Checklist

Every item traces to Anthropic's engineering guidance on building effective agents and on writing tools for agents, or to Google Cloud's production agent guidance. All checked 1 August 2026.

Work through it before an agent runs unattended, not after.

1. Data Access

- ☐ List every system the agent can read from.
- ☐ List every system it can write to. Keep these two lists separate on the page, because the difference is where the risk sits.
- ☐ For each write, name what happens if it writes the wrong thing.
- ☐ Confirm the agent cannot reach production data during testing. Anthropic recommends extensive testing in sandboxed environments.

2. The Tool Interface

Anthropic treats tool descriptions as prompt engineering in their own right, and says even small refinements to them can yield dramatic improvements.

- ☐ Every tool has a description that makes correct use obvious.
- ☐ Parameters and example usage are present, not just names.
- ☐ Arguments are shaped so mistakes are harder to make. Anthropic calls this applying poka-yoke principles.
- ☐ Error responses communicate specific and actionable improvements, rather than opaque error codes or tracebacks.
- ☐ Related tools are namespaced under common prefixes, to delineate boundaries between lots of tools.
- ☐ Long outputs use pagination, range selection, filtering or truncation, with sensible default parameter values.
- ☐ Formats avoid unnecessary overhead, and sit close to what the model has seen naturally occurring in text.

3. Evaluation

- ☐ An evaluation exists. Anthropic states that building one allows you to systematically measure the performance of your tools.
- ☐ Many example inputs have been run to see what mistakes the model makes, and the tools iterated on the results.
- ☐ The evaluation includes malformed inputs, not only well formed ones.
- ☐ Someone other than the author has read the failure cases.

4. Stopping Conditions

Anthropic names stopping conditions, such as a maximum number of iterations, to maintain control.

- ☐ A maximum number of iterations is set, in code, not in a prompt.
- ☐ The cap is a written requirement with a figure, not an afterthought.
- ☐ What happens at the cap is defined: stop, escalate, or hand back.
- ☐ Someone is told when the cap is hit.
- ☐ Note that per iteration cost varies with which tools are called, so the cap bounds duration rather than spend.

5. Human In The Loop Points

Anthropic describes agents that pause for human feedback at checkpoints or when they hit a blocker, and says human review remains crucial even when automated testing passes.

- ☐ Every action the agent can take is written down as a list.
- ☐ Actions that move money are marked.
- ☐ Actions that contact a customer are marked.
- ☐ Actions that delete anything are marked.
- ☐ The marked ones have a named person, not a named team.
- ☐ The list has been read by whoever owns the process, not only by whoever built it.
- ☐ Consider a delegated approval pattern. Google Cloud describes workflows where agents pause for human review while consuming zero compute resources.

6. Failure Handling

Anthropic notes the autonomous nature of agents means higher costs and the potential for compounding errors.

- ☐ Behaviour on a tool timeout is defined.
- ☐ Behaviour on a malformed tool response is defined.
- ☐ Retries are bounded, and a retry that keeps failing stops.
- ☐ The agent's planning steps are visible. Anthropic recommends prioritising transparency by explicitly showing them.
- ☐ You can reconstruct which step went wrong three weeks later.
- ☐ Consider checkpoint and resume. Google Cloud's patterns cover recovering from failures this way.

7. Before It Runs Unattended

- ☐ All six sections above completed, not skimmed.
- ☐ The iteration cap tested by deliberately hitting it.
- ☐ At least one checkpoint tested by refusing the action.
- ☐ The evaluation re-run after the last change, not before it.
- ☐ A named owner for when it misbehaves.

8. Sources

1. building effective agents
anthropic.com/engineering
2. writing tools for agents
anthropic.com/engineering
3. five guides to building and scaling production ready AI agents
cloud.google.com

All checked 1 August 2026.

Free Scoping Audit

If an agent mostly works and you cannot tell why it sometimes does not, the form asks for your website or store URL and a note on what is eating the time. The scoping audit costs nothing.

edm-uk.com/#form-start

Expert Digital Marketing Ltd. Accompanies the article at
edm-uk.com/agent-ai/build-ai-agents-production

Free audit:
edm-uk.com/#form-start