

Agent Observability Readiness Review

Use this checklist in your next design review. Tick each statement that is true for your agent today. Thresholds and intervals in the guide are recommended starting points; calibrate to your environment.

TRACE DESIGN

- ☐ Every run has one identifier that follows it through every step, tool and agent.
- ☐ Every span records the prompt, model and tool versions.
- ☐ The assembled context is recorded for each step.
- ☐ Policy decisions and human approvals are recorded for every action that changes data.

SIGNALS AND ALERTS

- ☐ Quality, reliability, cost and speed, safety, and human oversight signals are all tracked.
- ☐ Every alert has a named owner and a defined action.
- ☐ Online eval scores are tracked next to operational metrics.

DATA PROTECTION

- ☐ Personal data in traces is redacted or stored as a reference.
- ☐ Access to traces is role-based, and who viewed a trace is logged.
- ☐ Retention periods are agreed with legal and security for each risk level.

REVIEW AND IMPROVEMENT

- ☐ Sampled and flagged traces are reviewed on a regular schedule.
- ☐ Every finding becomes a fix or a new eval case.
- ☐ Changes to prompts, models and tools are logged and linked to traces.

Score: ____ of 13. Any unticked box is a gap to close before the agent handles real work.

Full guide: readyforrole.com/blog/observability-and-tracing.html | Series hub:
readyforrole.com/blog/enterprise-ai-agent-engineering.html | support@readyforrole.com