

Observability Maturity Assessment

For AI-Enabled and Cloud-Connected Systems

HOW TO USE THIS ASSESSMENT

What this tells you

Observability is not a dashboard problem. It is an operational trust problem. This assessment evaluates your organization's current observability maturity across five domains that matter most when AI-enabled systems are in production — or approaching production.

Score each question honestly. At the end calculate your total and review the maturity level that matches your score. The result tells you where your real risk is concentrated before you scale AI further.

SCORING KEY

0 — Not in place

1 — Partially in place

2 — Fully in place

DOMAIN 1 — TELEMETRY COVERAGE

Q1 · Telemetry Coverage

Do your AI-enabled systems emit structured logs, metrics, and traces to a centralized platform?

0	1	2
---	---	---

0 = No centralized logging · 1 = Partial coverage · 2 = Full telemetry across all AI workloads

Q2 · Telemetry Coverage

Are your cloud workloads — AWS and/or Azure — instrumented with telemetry that flows into a central observability stack?

0	1	2
---	---	---

0 = No cloud telemetry · 1 = Some workloads instrumented · 2 = All cloud workloads covered

Q3 · Telemetry Coverage

Is AI model inference — prompts, outputs, latency, token usage — logged and retained with defined policies?

0	1	2
---	---	---

0 = No inference logging · 1 = Basic logging exists · 2 = Full inference audit trail with retention policy

DOMAIN 2 — ALERTING AND THRESHOLDS

Q4 · Alerting

Have you defined SLOs and alerting thresholds for your business-critical systems?

0	1	2
---	---	---

0 = No defined thresholds · 1 = Informal thresholds exist · 2 = Formal SLOs with automated alerting

Q5 · Alerting

Does your team learn about system failures from monitoring tools — or from end users?

0	1	2
---	---	---

0 = Always from users · 1 = Sometimes from monitoring · 2 = Always from monitoring before users notice

Q6 · Alerting

Do you have alerts specifically for AI system degradation — drift, output quality changes, latency spikes?

0	1	2
---	---	---

0 = No AI-specific alerts · 1 = Some basic alerts · 2 = Comprehensive AI health monitoring

DOMAIN 3 — DEPENDENCY AND FAILURE VISIBILITY

Q7 · Dependencies

Is there a documented map of dependencies for your AI-enabled systems — data sources, APIs, integrations?

0	1	2
---	---	---

0 = No dependency map · 1 = Partial documentation · 2 = Full dependency map maintained and current

Q8 · Dependencies

Can your team trace a production failure to its root cause within one hour without tribal knowledge?

0	1	2
---	---	---

0 = No — requires key individuals · 1 = Sometimes · 2 = Yes — documented runbooks exist

Q9 · Dependencies

Are failure modes for your AI systems identified, documented, and tested before production deployment?

0	1	2
---	---	---

0 = No failure mode analysis · 1 = Partial · 2 = Failure modes documented and tested

DOMAIN 4 — OWNERSHIP AND ACCOUNTABILITY

Q10 · Ownership

Is there a named individual or function accountable for the operational health of each AI-enabled system?

0	1	2
---	---	---

0 = No clear owner · 1 = Informal ownership · 2 = Formal ownership assigned and documented

Q11 - Ownership

Does leadership have visibility into AI system health — or only find out when something fails?

0	1	2
---	---	---

0 = No leadership visibility · 1 = Ad hoc reporting · 2 = Regular operational health reporting to leadership

Q12 - Ownership

Is there a defined escalation path when an AI system produces unexpected or harmful outputs?

0	1	2
---	---	---

0 = No escalation path · 1 = Informal process · 2 = Formal escalation path documented and practiced

DOMAIN 5 — RESILIENCE AND RECOVERY

Q13 - Resilience

Have recovery time objectives been defined and tested for your AI-enabled systems?

0	1	2
---	---	---

0 = No RTOs defined · 1 = Defined but not tested · 2 = Defined, tested, and validated

Q14 - Resilience

Do your AI systems have graceful degradation — a defined behavior when a dependency fails?

0	1	2
---	---	---

0 = No graceful degradation · 1 = Partial · 2 = Graceful degradation designed and implemented

Q15 - Resilience

Is there a disaster recovery plan that specifically covers AI workloads and their data dependencies?

0	1	2
---	---	---

0 = No DR plan for AI · 1 = General DR plan exists but does not cover AI · 2 = AI-specific DR plan tested

YOUR SCORE

Add up your scores from all 15 questions. Maximum score is 30.

Total Score

Domain 1 — Telemetry

Domain 2 — Alerting

What your score means

Review the maturity level that matches your total score. Each level includes a plain-language description and recommended next steps.

Score: 0 – 9

Level 1 — Reactive

Your environment is operating on assumption. Issues are discovered by end users before your team knows about them. AI-enabled systems in this environment will fail in ways that are difficult to diagnose and expensive to recover from.

What this typically looks like:

- IT learns about outages from end users or customers
- No centralized logging or telemetry
- Dependencies are undocumented and poorly understood
- No defined ownership for AI system health
- Recovery from failures relies entirely on tribal knowledge

Recommended next step:

- Do not scale AI in this environment without addressing observability fundamentals first
- Start with centralized logging and basic alerting — these are the highest impact changes
- Consider an Operational Risk Diagnostic to understand the full scope of risk before proceeding

Score: 10 – 17

Level 2 — Developing

You have some visibility but significant gaps remain. Your team can detect some issues proactively but relies on manual processes and institutional knowledge for most operational decisions. Scaling AI in this environment will surface the gaps quickly.

What this typically looks like:

- Some monitoring exists but coverage is inconsistent
- Alerting is present but thresholds are informal or outdated
- Dependency mapping is partial — some systems are documented
- Ownership exists informally but is not formally assigned
- Recovery processes work but depend on specific individuals

Recommended next step:

- Prioritize closing the gaps in your highest-risk domains first
- Define SLOs for your most critical AI-enabled systems
- Document failure modes and escalation paths before expanding AI scope
- An observability implementation engagement can systematically close these gaps

Score: 18 – 24

Level 3 — Established

Your observability posture is solid. You have meaningful visibility into system health, defined ownership, and documented processes. You can scale AI with reasonable confidence — with some targeted improvements in your weaker domains.

What this typically looks like:

- Centralized logging and monitoring in place
- SLOs defined with automated alerting
- Dependency maps exist and are reasonably current
- Clear ownership and escalation paths documented
- Recovery processes tested at least periodically

Recommended next step:

- Focus on the domains where you scored lowest
- Add AI-specific observability — drift detection, inference logging, output quality monitoring
- Consider ongoing advisory to maintain and improve your observability posture as AI scales

Score: 25 – 30

Level 4 — Advanced

Your organization has strong observability discipline. You have the foundation to scale AI with confidence. The focus shifts from building visibility to maintaining it and extending it as AI capabilities expand.

What this typically looks like:

- Full telemetry coverage including AI inference
- Automated alerting with well-defined SLOs
- Dependency maps current and actively maintained
- Formal ownership and governance for all AI systems
- DR plans tested and AI-specific recovery validated

Recommended next step:

- Focus on extending observability coverage to new AI capabilities as they are deployed
- Consider predictive observability — anomaly detection and drift monitoring at scale
- An ongoing advisory retainer can help maintain this posture as your AI footprint grows

WHAT TO DO WITH YOUR SCORE

Start a conversation

This assessment is a starting point — not a complete picture. The Noctua Logic Operational Risk Diagnostic goes deeper across infrastructure, security, governance, and observability in 10 business days. If your score revealed gaps you were not expecting, that is exactly the conversation worth having.

jim.arnold@noctualogic.com

513-409-1222

noctualogic.com

[Schedule an Introductory Call](#)

Noctua Logic, LLC

The Logic of Stability — Timed for your success.

© 2026 Noctua Logic, LLC. All rights reserved.

noctualogic.com