

The AI Observability Evaluation Guide

What to Look for When Monitoring AI Agents in Production



Contents

- AI Is in Production. Is Your Observability Ready? 3
- What Is AI Observability? 3
- AI Observability Requires Enterprise-wide Visibility 4
- Key Criteria for Evaluating AI Observability Solutions 5
 - 1. Visibility Into Enterprise AI Usage and Adoption 5
 - 2. Diagnosing AI Performance Across the IT Environment.....6
 - 3. Understanding AI Usage and Cost 7
 - 4. Supporting Security, Governance, and Oversight.....8
 - 5. Integrate With Existing IT Operations 9
- The AI Observability Evaluation Checklist 10
- Choosing the Right AI Observability Solution 11
- See AI Observability in Action 11

AI Is in Production. Is Your Observability Ready?

AI is becoming part of everyday business operations. Organizations are deploying AI-powered applications, integrating large language models into business workflows, and introducing autonomous agents that interact with applications, data, and services.

For IT, this creates a new operational challenge.

AI-powered services depend on a complex combination of models, application code, APIs, databases, infrastructure, and networks. A performance issue in any one of these components can affect the entire AI experience. Meanwhile, AI adoption is expanding across the enterprise, introducing new questions about usage, costs, reliability, and governance.

Traditional monitoring tools provide important visibility into applications, infrastructure, networks and user experience. However, they don't currently

reveal what AI agents are doing, how AI services interact, or how problems propagate across the full environment.

IT needs to understand not only whether AI is running, but whether it's performing reliably and how it's impacting business operations and user experience. That requires a broader approach to AI observability.

This guide outlines what to look for when evaluating AI observability solutions and the questions to ask vendors to determine whether a solution can meet the needs of your IT environment.

What Is AI Observability?

AI observability is the ability to monitor, understand, and investigate the behavior, performance, and operational impact of AI applications and agents.

It combines AI-specific telemetry with information about the applications, infrastructure, networks, and services that support AI workloads.

Effective AI observability helps IT teams answer six fundamental questions:

1. What AI tools are being used across your enterprise?

Discover sanctioned and unsanctioned or Shadow AI, identify who's using which tools, and understand adoption patterns.

2. What are AI agents doing?

Monitor agent execution, interactions, and workflow performance.

3. Why is AI performing poorly?

Diagnose issues across applications, infrastructure, and networks.

4. What are our AI costs?

Understand consumption, spending, and identify opportunities to optimize spend.

5. Is AI operating within established controls?

Support governance, investigations, and operational oversight.

6. How does AI fit into existing IT operations?

Connect AI monitoring with established troubleshooting, incident management, and remediation workflows.

AI Observability Requires Enterprise-wide Visibility

Monitoring AI agents is only part of the challenge.

As AI becomes embedded in business operations, IT needs to understand how it's being used, how it's performing, and what impact it's having on the broader technology environment. That requires visibility across AI applications, infrastructure, and networks, along with insight into consumption, costs, and operational risks.

AI observability solutions vary in scope. Some focus on AI adoption and consumption, while others monitor agent behavior, application performance, or infrastructure. Each provides valuable insights, but without connecting these perspectives, IT teams may

struggle to understand what's causing problems, how issues affect users and business operations, and where to focus their efforts.

When evaluating AI observability solutions, look beyond individual monitoring capabilities. Consider how well a solution connects AI activity with the broader operational environment and helps IT teams identify, investigate, and resolve problems.

The following five evaluation areas provide a framework for assessing solutions against your organization's requirements.



Key Criteria for Evaluating AI Observability Solutions

The right AI observability solution depends on your organization's environment, priorities, and operational requirements. Use these five criteria to assess vendor capabilities, identify potential gaps, and determine which solutions best meet your needs.

1. Visibility Into Enterprise AI Usage and Adoption

Do you know who is using AI and which tools they are using?

AI adoption often extends well beyond centrally managed initiatives.

Employees use public AI assistants, business applications introduce embedded AI functionality, and developers integrate AI models and services into new applications. Some usage is approved and monitored. Other activity, often called Shadow AI, happens outside established IT oversight.

Understanding AI adoption requires more than knowing which tools are being used. Organizations need to see who's using AI, how frequently, and how adoption varies across departments and teams. These insights can reveal early adopters, identify groups that may benefit from additional training, and help organizations measure the impact of AI initiatives.

With a clearer picture of AI usage and adoption, IT and business leaders can encourage effective adoption, anticipate infrastructure demands, and make more informed decisions about AI investments.

What to Look For

An effective solution should help IT teams:

- **Discover AI activity:** Identify sanctioned and Shadow AI applications and services being used across the enterprise.
- **Understand who's using AI:** See which users, devices, and applications are accessing AI services.

- **Measure AI usage:** Monitor activity levels, traffic volumes, and frequency of use across AI tools and services.
- **Track adoption trends:** Monitor how AI usage evolves over time and assess the impact of training, enablement, and adoption initiatives.
- **Measure AI adoption:** Compare usage across departments and teams to identify early adopters, underutilized tools, and opportunities for additional training.
- **Support governance:** Identify unauthorized AI to reduce security risks and ensure compliance with organizational policies.

Questions to Ask Vendors

- Can you discover AI applications and services that IT has not explicitly configured for monitoring?
- Can you identify which users and devices are accessing specific AI tools?
- Can you track AI adoption and usage trends over time?
- Can you compare AI adoption and usage across departments, teams, and user groups?
- Can you identify adoption gaps and track whether training and enablement initiatives increase AI usage?
- Can you distinguish sanctioned AI services from Shadow AI?

What separates stronger solutions: The ability to discover AI services across the enterprise and provide meaningful insight into who's using them, how they're being used, and how adoption is evolving.

2. Diagnosing AI Performance Across the IT Environment

Can You Identify What's Affecting AI Performance?

AI applications and agents depend on an interconnected technology environment. A single AI-powered service may involve multiple model calls, tool invocations, APIs, databases, cloud infrastructure, GPUs, and enterprise networks.

When performance degrades or an AI workflow fails, the cause isn't always obvious. An agent may execute unnecessary steps, a model may respond slowly, infrastructure resources may become constrained, or network congestion may disrupt communication between services. Each can affect performance, reliability, and ultimately the user experience.

Understanding these issues requires visibility into both AI agent behavior and the broader IT environment. Monitoring individual components independently can leave IT teams struggling to connect the evidence needed to identify root causes and understand the impact on business operations.

Effective AI observability brings these perspectives together, helping IT teams diagnose problems, accelerate resolution, and anticipate capacity requirements as AI adoption grows.

What to Look For

An effective solution should help IT teams:

- **Monitor agent execution:** Trace agent workflows, model calls, and tool interactions to identify failures, unexpected behavior, and inefficient execution patterns.
- **Monitor across technology domains:** Correlate AI, application, infrastructure, network, and digital experience telemetry to understand end-to-end performance.
- **Understand service dependencies:** Identify relationships between AI services and the applications, infrastructure, and networks supporting them.

- **Diagnose performance issues:** Identify application bottlenecks, infrastructure constraints, network latency, packet loss, and other issues affecting AI services.
- **Accelerate root-cause analysis:** Use contextual intelligence and AI-assisted investigation to identify likely causes and guide remediation.
- **Assess user impact:** Understand how AI performance issues affect digital experience and business operations.
- **Plan for growth:** Analyze AI traffic patterns, resource utilization, and trends to anticipate infrastructure and network capacity requirements.
- **Validate resolution:** Confirm that service performance has recovered following remediation.

Questions to Ask Vendors

- Can you trace AI agent workflows and identify failed model calls, tool invocations, or unexpected execution patterns?
- Can you correlate AI agent activity with application, infrastructure, network, and digital experience telemetry?
- Can you identify dependencies between AI services and the underlying IT environment?
- Can you distinguish between AI workflow failures, application bottlenecks, infrastructure constraints, and network performance issues?
- Can you identify how AI performance issues affect end users and business services?
- Can you analyze AI traffic growth and resource utilization to anticipate capacity requirements?
- Can operators use conversational AI to investigate performance issues and identify likely root causes?
- Can you validate that remediation restored expected service performance?
- **What separates stronger solutions:** The ability to connect AI agent behavior and performance with full-stack operational context, enabling IT teams to identify root causes, understand business impact, and resolve issues faster.

3. Understanding AI Usage and Cost

Do You Know What Your AI Is Costing?

AI spending can grow quickly and unpredictably.

Different models, providers, applications, and workloads have different consumption patterns and pricing structures. Token usage, repeated model calls, inefficient workflows, and agent retries can contribute to unexpected costs.

Basic spending reports show what was consumed, but not necessarily why consumption increased, who is driving it, or whether the spend is delivering value.

More advanced cost analytics correlate consumption with usage patterns, users, teams, services, and outcomes to reveal the factors behind spending increases. This helps organizations identify cost drivers, uncover unused or underutilized investments, attribute costs accurately, and make more informed decisions about where to optimize or expand investment

IT teams need visibility that connects AI usage with operational context.

What to Look For

An effective solution should help IT teams:

- **Track consumption:** Monitor token, credit, or usage-based charges and other relevant consumption metrics.
- **Understand spending:** Analyze AI costs across applications, models, agents, and services.
- **Attribute usage:** Identify the teams, users, or applications driving AI consumption and spending.

- **Identify anomalies:** Detect unexpected spikes or changes in AI usage and costs.
- **Uncover overspending:** Identify underutilized tools, redundant subscriptions, and opportunities to reduce unnecessary spending.
- **Optimize investments:** Compare usage and costs across AI tools and services to identify opportunities to consolidate or rebalance spending.

Questions to Ask Vendors

- Can you attribute AI usage and costs to specific users, teams, applications, or services?
- Can you track spending trends and identify unexpected increases in AI consumption or costs?
- Can you compare usage and spending across different AI tools, models, and providers?
- Can you identify underutilized AI tools, redundant subscriptions, and opportunities to reduce unnecessary spending?
- Can you identify opportunities to consolidate AI tools or rebalance spending based on actual usage?
- Can you alert teams when consumption or spending exceeds defined thresholds?
- How do you reconcile estimated AI costs with actual provider billing

What separates stronger solutions: The ability to explain what's driving AI consumption and identify opportunities to improve efficiency, not simply report spending totals.

4. Supporting Security, Governance, and Oversight

Can You Maintain Oversight as AI Adoption Grows?

As AI becomes embedded in business operations, organizations need greater visibility into where AI is being used and how AI-powered applications and agents behave in production.

Shadow AI may operate outside established policies. AI applications and agents may depend on sensitive systems, data, and external services. And as adoption expands, organizations need a clearer understanding of where AI is running, how it is performing, and where operational or compliance risks may emerge.

AI observability does not replace dedicated AI security, compliance, or governance solutions. Instead, it provides operational visibility and evidence that can help organizations identify risk, investigate issues, and make more informed governance decision.

What to Look For

An effective solution should help IT teams:

- **Identify shadow AI:** Reveal AI services being used outside established approval processes.
- **Understand activity:** Provide relevant records of AI application and agent execution.
- **Support investigations:** Help teams reconstruct events and investigate unexpected behavior. It should provide visibility into anomaly blast radius, root cause, and time to resolution.

- **Protect observability data:** Apply appropriate enterprise controls to the telemetry and operational data used for AI observability.
- **Identify potential risk:** Surface usage, performance, and operational conditions that may warrant further security, compliance, or governance review.
- **Support auditability:** Retain relevant operational evidence according to organizational requirements.

Questions to Ask Vendors

- Can you identify unsanctioned AI usage across the enterprise?
- What AI agent activities are captured and retained?
- Can you reconstruct agent workflows during incident investigations?
- How do you protect sensitive information in transit and at rest?
- Does the solution support role-based access controls?
- Can retention policies be configured to meet organizational requirements?

What separates stronger solutions: Operational transparency that supports governance without overstating monitoring capabilities as security enforcement.

5. Integrate With Existing IT Operations

Will AI Observability Fit Into How Your IT Teams Work?

AI applications depend on the same networks, infrastructure, applications, and services that IT teams already manage. When an AI service fails, troubleshooting may require collaboration across application development, NetOps, cloud operations, infrastructure, and service management teams.

Adding another isolated monitoring tool can make investigations more difficult. IT needs a way to incorporate AI observability into existing operational workflows, making it easier to connect data across systems, uncover insights, and accelerate problem resolution.

What to Look For

An effective solution should help IT teams:

- **Connect operational data:** Correlate AI telemetry with existing application, infrastructure, network, and digital experience monitoring.
- **Integrate with ITSM:** Support established incident creation, escalation, and tracking processes.
- **Enable conversational investigation:** Use LLM-powered and agentic AI interfaces to explore observability data, uncover relationships, identify likely root causes, and guide resolution using natural language.
- **Guide remediation:** Provide recommendations and connect with established automation workflows.

- **Support collaboration:** Help multiple IT teams investigate problems using shared operational context.
- **Protect existing investments:** Extend the value of deployed monitoring tools rather than requiring unnecessary replacement.

Questions to Ask Vendors

- Can you correlate AI-specific telemetry with existing network, infrastructure, and digital experience data?
- Can AI incidents be created and managed through our ITSM platform?
- Can IT teams use a conversational AI interface to query and analyze observability data across multiple technology domains?
- Can agentic AI autonomously investigate issues, correlate evidence, identify likely root causes, and recommend remediation?
- Does the solution provide AI-assisted diagnosis and recommended and prioritized remediations?
- Can remediation workflows incorporate human approval and existing automation tools?
- What instrumentation, infrastructure, and operational overhead are required?

What separates stronger solutions: The ability to combine AI observability with existing operational data, conversational investigation, and agentic AI to accelerate diagnosis, guide remediation, and support more autonomous IT operations.

The AI Observability Evaluation Checklist

Use this checklist to evaluate vendors against your organization's requirements. For each criterion, assign a score based on demonstrated capability:

1: Not supported

2: Limited support

3: Partially supported

4: Supported

1. Enterprise AI Adoption

- ☐ Discovers sanctioned and Shadow AI services
- ☐ Identifies AI usage by users, devices, and applications
- ☐ Tracks adoption trends and usage patterns
- ☐ Provides visibility across hybrid and remote environments

2. AI Agents and Workflows

- ☐ Traces multistep agent workflows
- ☐ Monitors execution times, failures, and retries
- ☐ Correlates agent behavior with application dependencies
- ☐ Identifies abnormal or inefficient execution patterns

3. Full-Stack Performance

- ☐ Correlates AI, application, infrastructure, and network telemetry
- ☐ Distinguishes model, application, infrastructure, and network delays
- ☐ Identifies dependencies across AI-powered services
- ☐ Provides AI-assisted root-cause analysis
- ☐ Supports conversational troubleshooting
- ☐ Validates performance following remediation

4. AI Usage and Cost

- ☐ Tracks model usage and token/credit consumption
- ☐ Attributes consumption and costs to workloads or teams
- ☐ Detects unusual spending and consumption patterns
- ☐ Supports cost optimization decisions

5. Governance and Oversight

- ☐ Identifies unsanctioned AI activity
- ☐ Provides records of agent execution
- ☐ Supports operational investigations and auditability
- ☐ Protects sensitive telemetry through access controls and encryption
- ☐ Integrates with security and governance processes

6. IT Operations Integration

- ☐ Integrates with existing observability tools
- ☐ Connects AI telemetry with broader operational context
- ☐ Integrates with ITSM workflows
- ☐ Provides AI-assisted investigation and guidance
- ☐ Supports established remediation and approval workflows

Choosing the Right AI Observability Solution

As organizations put more AI-powered applications and agents into production, IT needs visibility into how they are being used, how they are performing, and what happens when something goes wrong.

That visibility needs to extend beyond the AI itself. IT teams need visibility into how AI interacts with applications, infrastructure, and networks. They also need to understand how problems affect users and business services.

The right solution should help teams find and troubleshoot problems, understand usage and costs, and work with the tools and processes they already use to manage IT.

When evaluating vendors, look beyond the number of metrics, dashboards, and AI models they support. Ask whether the solution gives IT the visibility and context to keep AI services performing reliably as their use grows.

See AI Observability in Action

Riverbed AI Assurance gives IT visibility into how AI is being adopted, what it costs, and how AI-powered applications and agents are performing in production. It brings AI activity together with broader operational context to help teams understand problems, troubleshoot faster, and manage AI as its use grows.

See how Riverbed AI Assurance puts AI observability into practice.
riverbed.com/products/ai-assurance/.



About Riverbed

Riverbed, the leader in AI observability, helps organizations optimize their users' experiences by leveraging AI automation for the prevention, identification, and resolution of IT issues. With over 20 years of experience in data collection and AI and machine learning, Riverbed's open and AI-powered observability platform and solutions optimize digital experiences and greatly improve IT efficiency. Riverbed also offers industry-leading Acceleration solutions that provide fast, agile, secure acceleration of any app, over any network, to users anywhere. Together with our thousands of market-leading customers globally – including 95% of the FORTUNE 100 – we are empowering next-generation digital experiences. Learn more at riverbed.com.