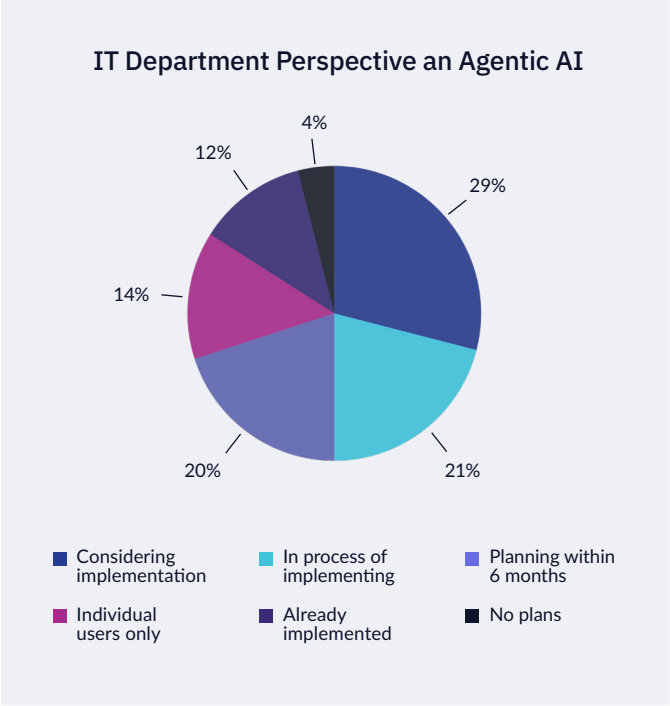
GUIDE

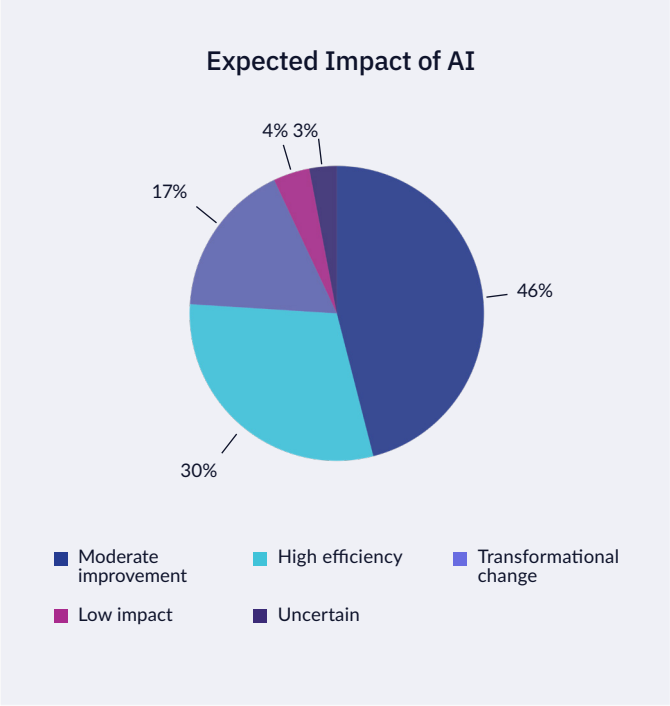
AI Fact or Fiction: 10 Questions to Ask MDR Providers About AI Capabilities

As cyber threats grow more advanced and response times become critical, Managed Detection and Response (MDR) providers are racing to integrate AI – specifically Agentic AI – into their platforms.

Agentic AI is a type of software that can perceive a task, independently plan and reason, and act autonomously with minimal human intervention. What's more, adoption of Agentic AI software is spreading – fast. In fact, Gartner **predicts** that by 2028, nearly one-third of all interactions with GenAI services will "use action models and autonomous agents for task completion".

For their part, most IT/Security leaders are adopting Agentic AI. New research by Georgian, a growth-stage investment firm, shows that **74% of IT respondents** have already started implementing, or are in the process of implementing, Agentic AI. Moreover, 30% anticipate high efficiency boosts while 46% expect moderate incremental improvements to existing workflows:





MDR vendors that are incorporating Agentic AI into their services promise to accelerate detection, reduce alert fatigue, and close the skills gap by automating tasks traditionally handled by human SOC Analysts. These systems continuously adapt and learn from real-time threats, providing proactive defense and automated remediation to ensure threats are contained before they disrupt business operations.

But not all AI platforms are created equal. While some MDR providers are deploying production-grade Agentic AI that actively drives threat investigations and response at scale, others are still in beta with little operational maturity.

For security leaders, the challenge isn't just identifying which vendors have AI – it's determining which ones are truly using it to deliver faster, more accurate, and more secure outcomes today.

In this guide, we outline key evaluation criteria to help you cut through the noise and assess whether an MDR provider's AI capabilities, including their claims about Agentic AI, are built to help you accelerate your security operations reliably, transparently, and at scale.

By the end, you'll be equipped to identify providers that deliver:

- **Operationally mature, outcome-driven AI** that's in full production, not just in beta
- **Human-in-the-loop approach with Agentic AI** that augments analyst-level decision-making, with expert validation embedded into every critical step of the threat detection and response process
- **Clearly defined autonomous actions** with built-in human oversight and role-based control
- **Resilience through adversarial testing** and secure deployment architecture
- **Tangible ROI** demonstrated through measurable improvements in speed, accuracy, and cost

These criteria will ensure alignment with your organization's specific security requirements and risk tolerance, enabling you to select a provider that suits your specific business needs.

Choosing an MDR Provider with AI Capabilities

Given the critical role that MDR providers play in your cybersecurity program, it's imperative to thoroughly assess their claims regarding their Agentic AI capabilities. The following evaluation criteria and questions will help you qualify providers, ensuring their AI technology effectively supports your organization's unique security requirements.

Is your AI system fully operational, or still in beta? What is your roadmap, and how is AI being applied across detection and investigation today?

It's essential to understand whether the MDR provider you're evaluating has AI capabilities that are production-ready or still under development. Many vendors are in early beta or pilot phases but present their AI as fully mature. Ask for clear evidence that the AI is active in real-world environments, driving measurable outcomes, and not just part of a roadmap slide.

Go deeper by asking how the AI is being applied today. Is it only involved in initial threat detection, or is it actively contributing to investigation and response? In other words, is their AI capable of enriching alerts, correlating indicators, and generating incident narratives that mirror the work of a skilled analyst? MDR vendors operating at scale should be able to show how their AI supports investigation quality and reduces time-to-resolution across thousands of customer environments.

Lastly, ask for visibility into their AI roadmap. If parts of the system are still evolving, what are the near-term development milestones? This helps clarify whether the provider has a real long-term vision or if they're simply responding to market pressure with partial solutions.

What actions can your Agentic AI take autonomously, which require human approval, and how are these limitations visible and controllable from anywhere (e.g., mobile app)?

Understanding which actions your MDR provider's AI agent can take independently (versus those requiring human intervention) is critical for ensuring appropriate control over incident response. Look for providers that explicitly document autonomous capabilities that specify exact actions AI can perform independently, such as endpoint isolation, file quarantine, or user account resets. These should be clearly defined, tested actions tied to specific threat scenarios.

More importantly, you should know how the MDR provider limits AI autonomy through role-based approval workflows for high-impact decisions. These workflows should ensure that any actions affecting business continuity, such as disabling accounts or altering configurations, require human validation, with escalation protocols that support after-hours decision-making. In other words, humans should always be in the loop. Also, ask if you'll get mobile-ready override features for rapid response in time-sensitive situations.

A complete audit trail must accompany every action the AI takes. Leading platforms will log both autonomous and human-led interventions and make this data easily accessible through the platform or mobile interface. This level of visibility helps ensure the AI operates within set boundaries and provides the accountability required for compliance and internal oversight.

3 How do you ensure the AI's decision-making process is transparent and auditable, suitable for regulatory compliance, forensic investigations, and board reporting?

IT/Security teams need full insight into how Agentic AI makes decisions, especially when those decisions inform or initiate threat response. A mature MDR provider should offer detailed, step-by-step reasoning behind each AI-driven action, avoiding any "black-box" operations. This includes an evidence trail that shows not only what action was taken, but why it was taken, and what indicators or context contributed to the decision.

Daily operational summaries are a good bonus feature to help teams translate AI activity into measurable value. These should highlight which threats were investigated, what mitigation steps were triggered, and where AI improved efficiency by reducing manual effort. This supports internal reporting needs and equips security leaders to demonstrate ongoing program performance to executive stakeholders.

Providers should also offer exportable evidence packages in formats such as JSON for use in forensic investigations or audits. Ensure that the MDR provider maintains clear retention policies that align with regulatory standards like PCI DSS, HIPAA, or GDPR, enabling retrospective analysis of historical decisions.

n you describe your adversarial testing program against threats like prompt-injection, model drift, and data-poisoning attacks on your Agentic AI?

AI systems integrated into security workflows must be continuously validated against adversarial threats that target their logic, behavior, and inputs. So, you should ask providers to detail how they conduct adversarial testing, including how often they simulate attacks such as prompt injection, data poisoning, and model drift. These tests should challenge the AI's ability to make accurate decisions under manipulated or degraded input conditions.

Look for a formal program that includes red-team exercises and scenario-based evaluations focused specifically on the AI layer. The goal is to identify weaknesses in model decisioning and assess whether the system maintains integrity under stress.

In addition, confirm that the MDR provider applies secure model management practices. This includes controlled update processes, embedded guardrails to prevent drift, and visibility into recent patching and tuning activity. Availability of third-party testing reports or summaries of remediation activity further demonstrates that the provider treats AI resilience as a core part of their security posture.

Which AI models, data sources, and upstream services underpin your AI system, and how do you handle their disruption, deprecation, or breach?

Understanding the components that power your MDR provider's AI is essential for evaluating system reliability and operational risk. Ask for a software bill of materials (SBOM) specific to their AI stack – an AI-SBOM, so to speak. This should identify the models in use, third-party data sources, upstream services, and the roles these components play in detection and response processes.

Just as critical is how the provider handles service degradation or disruption. A mature provider will have contingency plans in place for model failure, data source outages, or vendor breaches. This may include fallback models, alternative enrichment pipelines, or deployment flexibility such as sovereign cloud or on-premises options to maintain operational continuity.

In certain high-control or regulated environments, you may also require the ability to temporarily disable AI-driven decision-making and fall back to manual, analyst-led operations. Confirm whether the MDR provider supports human-only workflows as an emergency or policy-driven control. This ensures that when confidence in upstream services is in question (or during internal policy audits), your organization retains full control over threat response actions.

You should also understand the contractual obligations tied to third-party AI sub-processors. Providers should be transparent about their notification processes and offer clear escalation paths in the event of supply chain or model-related incidents that could impact your threat coverage or data security.

How do you quantify and communicate the effectiveness and ROI of your Agentic AI beyond anecdotal claims?

AI performance claims don't mean anything without measurable outcomes. The MDR provider should deliver comparative data showing how their AI capabilities improve your KPIs, including Mean Time to Respond (MTTR), alert volume reduction, false-positive suppression, and analyst touch time. Ideally, this should include pre- and post-deployment benchmarks.

Ask for examples that illustrate how the AI supports Tier 1 triage, enriches context for investigations, or shortens containment timelines. Some providers can demonstrate high analyst alignment, such as 95% agreement with Tier 3-level conclusions, and initial host isolation rates exceeding 99%, which directly translates to reduced lateral movement risk.

To validate these claims, request data-backed case studies that show how the AI impacted cost savings, improved detection precision, or supported operational scale. The ROI of Agentic AI should be evident in both tactical gains and strategic outcomes, not just anecdotal success stories.

How do you manage continuous learning without compromising data privacy, ensuring that my organization's data doesn't leak into multi-tenant training environments?

Continuous learning is essential to keeping AI systems effective, but it must be done without exposing sensitive customer data. MDR providers should be able to demonstrate how your telemetry is logically and physically segmented to prevent cross-contamination with other clients' data or public training models.

Ask whether the provider supports techniques like synthetic gradient updates, private fine-tuning enclaves, or federated learning that preserve privacy while still allowing the model to evolve. Additionally, data usage policies should be clearly defined, including opt-in or opt-out mechanisms for any use of your data beyond the operational scope.

Transparency in how data is handled—and guarantees around non-use in shared or multi-tenant training—are especially important for regulated industries and global enterprises with strict data sovereignty requirements.

8 How does your solution enable collaborative response between our internal teams and your analysts?

Effective incident response depends on seamless coordination between your internal IT/Security team and your MDR provider's own SOC Analysts. Look for platforms that support continuous collaboration through a centralized portal for shared notes, case timelines, and shared notes. These features ensure that context is maintained even as teams or shifts change, eliminating gaps and unnecessary confusion between parties.

In high-pressure scenarios, real-time collaboration tools like in-platform chat, threaded comments, and update notifications are critical. They reduce response delays, minimize miscommunication, and support consistent handoffs across stakeholders. Role-based tagging and assignment capabilities further help clarify accountability during incident handling.

A collaborative MDR model doesn't just streamline response; it ensures that all actions are traceable, auditable, and contextually aligned across both the customer and provider sides.

9 How does your solution balance AI-driven automation with expert human oversight, and how is that human involvement embedded into your detection and response workflows?

While AI capabilities, including the use of AI agents, can accelerate threat detection and response, the most effective solutions integrate human expertise at key decision points. MDR providers should outline exactly where AI operates autonomously, such as endpoint isolation, and where actions require real human validation, especially for decisions that could affect business operations or sensitive systems.

For instance, human involvement should be embedded into threat triage workflows, threat hunting activities, and containment decisions. AI can assist by providing correlation, prioritization, and contextual enrichment, but final containment steps should follow a clear chain of human approval.

Best-in-class MDR providers use Analyst feedback to tune the AI, close logic gaps, and minimize false positives. So, ask whether the provider maintains feedback loops where Analysts can train or correct the AI's logic over time. This continuous tuning, guided by subject matter experts, is essential for reducing false positives and improving long-term system reliability.

10 What steps have you taken to ensure your AI implementation is secure and suitable for enterprise-scale, regulated environments?

AI systems used in cybersecurity must meet the same security and compliance standards as the environments they protect. Moreover, there's no doubt that AI regulations will continue to evolve globally so providers should clearly articulate their compliance preparation strategy and regulatory readiness plans.

Providers should be able to explain how their AI is deployed, including whether it operates in containerized environments monitored 24/7 and if it's governed by a formal oversight program that includes model validation, access controls, and data handling policies.

A secure implementation should also support audit readiness. That means maintaining full records of AI-led decisions, analyst interventions, and response actions—all tied to a documented reasoning process. These logs should be accessible during compliance reviews, internal audits, or post-incident investigations.

Ask about the provider's track record supporting regulated sectors and whether their AI deployments have been reviewed against frameworks like SOC 2, ISO 27001, or industry-specific compliance requirements. At enterprise scale, security architecture matters as much as AI capability.

Accelerate Your Security Program with AI You Can Trust, Outcomes You Can Prove

eSentire's **Atlas Expert AI** is fully embedded into our Atlas XDR platform and included as part of your MDR service. Designed to scale human expertise, not replace it, Atlas Expert AI gives your security operation a competitive edge by providing transparency, context and validation previously unattainable in minutes.

Our Difference	Your Results
Production-ready Agentic AI	Get immediate value from day one with a mature AI platform that delivers proven results at scale.
Tier 3 Analyst-level Precision and Role-based Oversight for Autonomous Actions	Trust that Atlas Expert AI can drive high-confidence investigations with 95% alignment to senior SOC analysts and automatically take low-stakes containment actions (e.g., endpoint isolation) while maintaining human approval workflows for high-impact decisions, reducing risk without giving up control.
First Host Threat Isolation	Contain threats before they spread; 99.3% of threats are isolated at the first point of compromise, drastically minimizing lateral movement and blast radius.
Multi-Agent Cognitive Architecture	If an orchestrated response isn't possible, our platform equips our SOC team with the insights they need to perform deep investigation and execute manual containment, delivering a Mean Time To Contain of 15 minutes.
Noise Reduction at Scale	Eliminate alert fatigue with 99.99% noise suppression, allowing your team to focus on true positives while we manage the rest.
35% Faster Threat Intelligence	Stay ahead of attackers with threat intel that outpaces standard commercial feeds by 35%, delivering faster detection and more timely protections.
Human-in-the-Loop Approach	Retain control and confidence with every AI-led action validated by expert analysts, ensuring decisions align with your policies and operational context.
24/7 Expert-Guided SOC Operations	Every AI-led decision is backed by a 24/7 SOC team with 96% analyst retention and an average tenure of 6 years, delivering consistent oversight at all times.

Ready to get started?

Reach out to connect with an eSentire security specialist and build a more resilient security operation today.

IF YOU'RE EXPERIENCING A SECURITY INCIDENT OR BREACH CONTACT US 1-866-579-2200

eSENTIRE