REPORT

STATE OF AI CYBER SECURITY 2024

Industry Perspectives on the Growing Role of AI in Cyber Security

DARKTRACE

Table of Contents

Executive Summary

Keeping up with the latest threats isn't easy for cyber security professionals. Akin to untying the legendary Gordian knot, stopping waves of unknown threats across every IT domain before one gets in is intricately complex. Recent advancements in artificial intelligence (AI)—what's been called the AI revolution—are upping the ante even further. Security leaders must find strategies to level the playing field and evolve defenses in tandem with the sophistication of attacks.

The Impact of AI on Threats and Defense

We recently surveyed nearly 1,800 security leaders and practitioners across a broad array of global industries. Our research was conducted to understand how the adoption of new AI is affecting the threats stakeholders face, how they are responding, and AI's role in prevention, threat detection, incident response, and recovery workflows.

AI's effects on the threat landscape are already being felt. A majority of survey participants (74%) report their organizations are seeing significant impacts from AI-powered cyber threats. An even greater majority (89%) believe that AI-powered threats will continue to trouble their organizations well into the future.

Many security leaders and practitioners feel they are not ready for what's to come. Two in every three survey participants (60%) believe their organizations are inadequately prepared to defend against AI-powered threats and attacks. AI may help reduce the industry-wide skills shortage that has held defenders back, but only if practitioners can maximize the value of these new tools in real-world workflows. To do so, they will need to know which AI-powered solutions will deliver on their promise, and which are simply hype.

Participants in our survey *do* understand that AI-powered security solutions are a must-have for defending against AI-powered threats. Nearly all respondents (96%) believe that AI-driven security solutions significantly improve the speed and efficiency of prevention, detection, response, and recovery. A large majority (71%) are also confident that AI-powered security solutions will be able to detect and block AI-powered threats.

However, there remains a significant need for education within the field. Many security stakeholders are not sure which types of AI are being applied in solutions. With generative AI and large language models (LLMs) getting so much press, there's a tendency to overestimate how often they are used in security tools. Leaders are confident that their organizations are effectively mitigating AI security risks, while practitioners have less faith in these efforts. Professionals with hands-on experience may be less likely to be swayed by all the generative AI hype. Instead, they're looking to improve integrations between their tools, reduce the cost and complexity of their cyber security stacks, and adopt a platform-oriented approach when implementing new tooling.

Evolving in Tandem with the AI Revolution

To effectively aid defenders, AI must:

- Personalize security for every organization, since risks depend on unique attributes like users, infrastructure, and assets
- Not interfere with operational workflows or employee productivity
- Be transparent about its decision-making and work alongside human teams to simplify time-consuming tasks like alert triage
- Integrate with the wider security stack seamlessly while helping humans make informed decisions based on comprehensive data
- Assist security teams at every stage of the event lifecycle, making recommendations and suggesting next steps

Getting the Most from AI-Driven Solutions

Cyber security leaders and practitioners must apply the right types of AI in the right places within their tool stack to recognize and neutralize threats at machine speed. They must leverage unsupervised machine learning (ML) algorithms that can continuously train themselves to understand what's normal in their organizations so they can quickly and automatically recognize deviations from baseline. These should be applied in conjunction with other AI methods such as supervised ML, LLMs, generative adversarial networks (GANs), graph theory (to reveal complex relationships), anomaly detection, and generative AI. Using the right mix of AI types helps accurately identify the threats that yesterday's tools would have missed. Most importantly, AI should be transparent, explainable, and privacy-preserving.

In the report that follows, you will learn more about what we discovered from the survey. We'll also share some recommendations for addressing today's top challenges.

Key Findings

74%

of survey respondents are seeing AI-powered cyber threats significantly impact their organizations.

96%

of survey participants believe AI-driven security solutions are a must-have for countering AI-powered threats, significantly improving the speed and efficiency of prevention, detection, response, and recovery.

60%

of respondents fear their organizations are not adequately prepared to defend against AI-powered threats and attacks.

79%

of **IT security executives** believe their organizations have taken steps to reduce the risks associated with using AI.

VS

54%

of **hands-on practitioners** believe their organizations have taken steps to reduce the risks associated with using AI.

15%

of security stakeholders think traditional solutions (those not using AI) can detect and block AI-powered threats.

31%

of respondents need cyber security vendors to explain what types of AI are being used in their solutions and why.

88%

of organizations prefer a platform approach over implementing individual point products.

INTRODUCTION

The AI Revolution and Cyber Risk

Recent advances in the development of AI are radically changing the very nature of cyber security.

As attackers and defenders race to harness the full potential of AI, they're engaging in a strategic competition to build, implement, and use new tools in innovative ways to gain an edge.

The Shifting Threat Landscape

With generative AI and LLMs now widely used, malicious actors are launching attacks at machine speed and scale. This compounds an already-pressing challenge faced by security teams, a cycle in which more complex computing environments increase the attack surface, creating a need for additional visibility. This, in turn, results in more alerts.

Threat researchers have already observed a notable increase in email-borne and novel social engineering attacks since ChatGPT was released, but this trend represents only the tip of the iceberg.

In the months and years ahead, we expect to see more:

- Novel phishing campaigns that are highly convincing
- Automated creation of malicious code and multi-stage attacks
- Deepfakes designed to elicit trust
- CAPTCHA-breaking AI techniques
- Generative AI employed in open-source intelligence-gathering
- Advanced reasoning and decision-making capabilities in autonomous agents

As these types of threats grow, the cyber security industry is hastening to integrate AI into technologies used across workflows in prevention, detection, response, and recovery. However, this will place new demands upon defenders—practitioners and leaders alike.

Decision-makers will be tasked with understanding the evolution of products and their features (which is not easy, given the complexity of AI and how recently it has come to the fore). Security analysts, incident responders, and architects will need to know how to work with this new technology on a hands-on basis. Not all AI-based solutions are equal, and not all of them can drive the risk reduction CISOs are hoping for.

Generative AI has attracted the most attention lately, but it is not the only type of AI relevant for cyber defense. It can serve in a limited number of security use cases. It's good at summarizing incidents and enabling analysts to interact with algorithms using easy-to-understand natural language prompts. And GenAI is fairly simple for vendors to integrate into existing solutions. But other forms of AI—such as supervised and unsupervised machine learning, large language models (LLMs), and generative adversarial networks (GANs) can do much more.

But they need to be understood to be useful.

This means that security leaders and practitioners will need a more nuanced grasp of AI: what it is, how it works, and how its full potential can best be achieved in real-world operational workflows. Building effective defenses requires education, mindset shifts, and new ways of thinking.

87% of respondents believe that AI-powered threats will have a significant impact on their organizations for months and years to come.

[1] Darktrace, "Major Upgrade to Darktrace/Email™ Product Defends Organizations Against Evolving Cyber Threat Landscape, Including Generative AI Business Email Compromise and Novel Social Engineering Attacks," Available at: https://darktrace.com/news/darktrace-email-defends-organizations-against-evolving-cyber-threat-landscape

A PERILOUS SHIFT

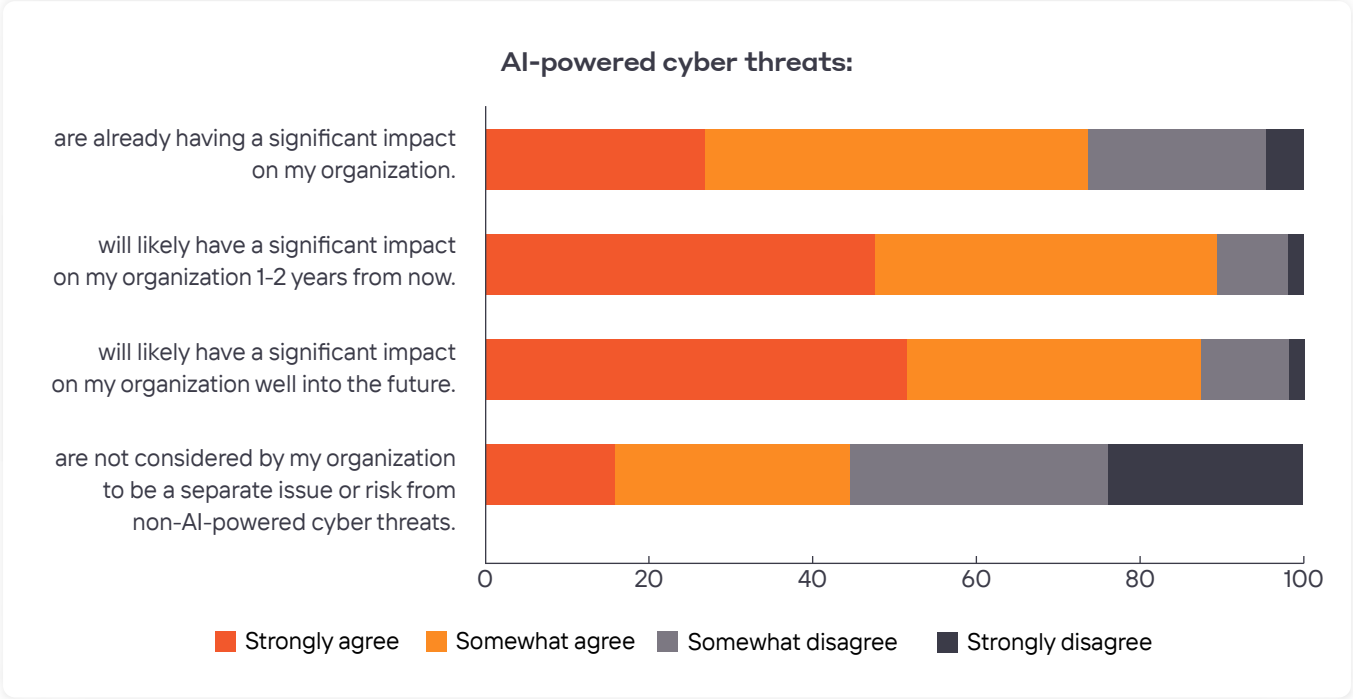
The Impact of AI on the 2024 Cyber Threat Landscape

Within the last year, as generative AI systems like ChatGPT and Google Gemini have made headlines and dominated conversations, attackers have set their sights on harnessing the power of this emerging technology.

Our threat researchers have observed significant growth in the breadth, scope, and complexity of threats that organizations are confronting.[2] While it is difficult to be certain exactly how much of this activity is directly attributable to the generative AI boom, we expected to see security leaders expressing concerns about a rise in AI-powered cyber threats. That is exactly what our survey found.

[2] Darktrace, End of Year Threat Report: Analysis of the Second Half of 2023, Available at: https://assets-global.website-files.com/626ff19cdd07d1258d49238d/65c10e74dd798acb4fc99c72_End_of_Year_Threat_Report_2023_Final.pdf

Nearly three in four organizations are already feeling the impact of AI-powered threats.



Participants report that AI is already having a major impact on the threats their organizations are confronting. Nearly three-quarters (74%) state AI-powered threats are now a significant issue. Almost nine in ten (89%) agree that AI-powered threats will remain a major challenge into the foreseeable future, not just for the next one to two years.

89% of survey participants agree that AI-powered threats will remain a major challenge well into the forseeable future.

These concerns didn't vary substantively across organizational sizes or by role. However, there were notable differences across regions:

- A large majority (84%) of respondents in Asia-Pacific are already feeling the impact of AI-powered threats.
- Only 71% of survey participants in Latin America agreed that AI-powered threats were already having a significant impact. Respondents from this region also expressed more optimism about the future.

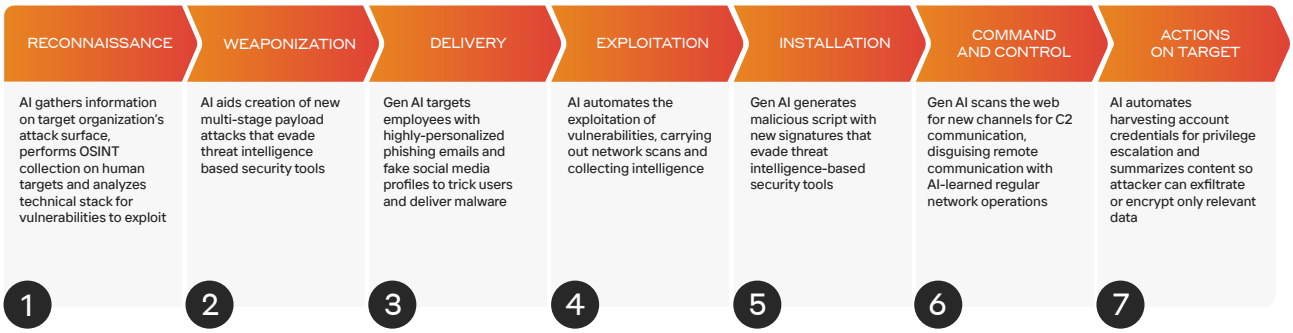
There was a near-even split between the participants whose organizations considered AI-powered threats to be substantively different from non-AI-powered threats and those that did not. Only a slight majority (56%) thought AI-powered threats were a separate issue.

AI has already accelerated the speed and scale at which some types of attacks can be launched. Various forms of AI enable attackers to automate the creation of unique malware scripts and multi-stage payloads. Generative AI can create large volumes of highly personalized, convincing phishing attacks. AI can repeatedly change the signatures and hashes associated with malware files, so that traditional threat intelligence-based detection tools (which use known threat attributes to identify what's malicious) will no longer be able to keep up.

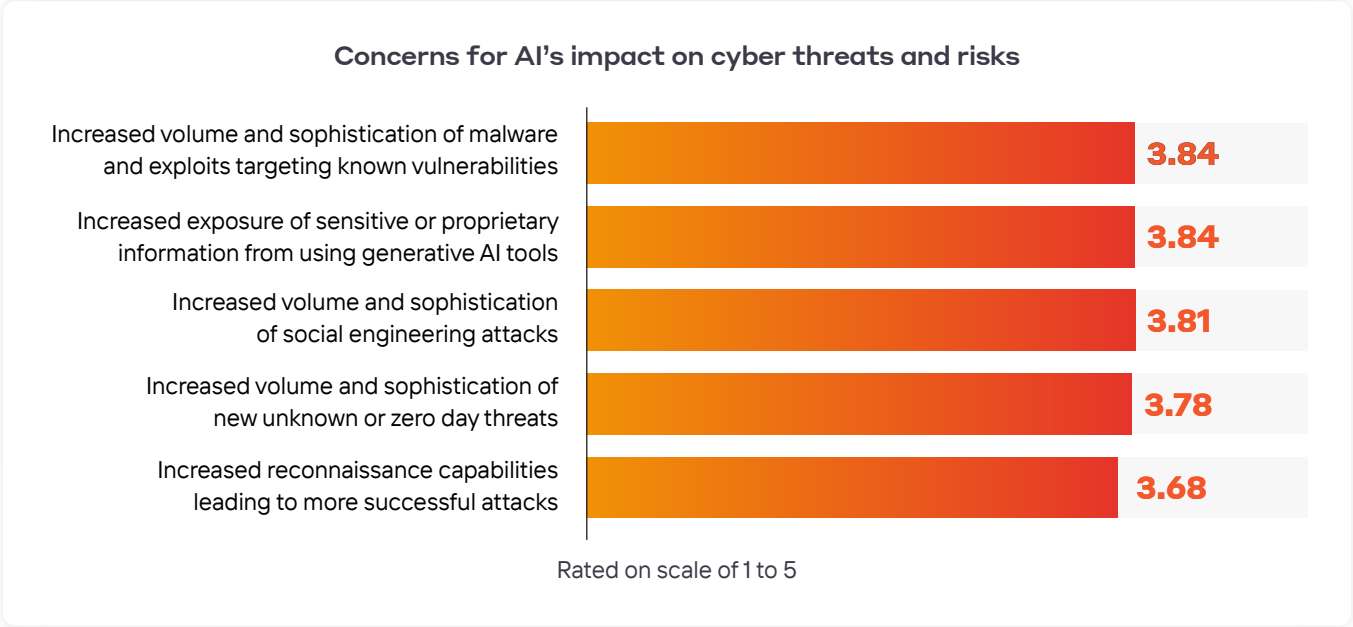
But there are few, if any, reliable methods to determine whether an attack is AI-powered. It's likely that AI will be applied to amplify attack effectiveness across every stage of the kill chain.

Identifying exactly when and where AI is being applied may not ever be possible. Instead, defenders will need to focus on preparing for a world where threats are unique and are coming faster than ever before.

A hypothetical cyber attack augmented by AI at every stage



Security stakeholders are worried about a broad array of AI-driven threats, including more sophisticated malware and increased numbers of social engineering attacks.



These results show that concern is high across all areas. Even the lowest-rated risk, "increased reconnaissance capabilities leading to more successful attacks," is still above the midpoint of the spectrum. This is reasonable, since each of these issues represents a real-world threat that is growing in prevalence and likely to become more of a concern in the future.

AI is having a far-reaching impact on both the external threat landscape and internal risks. Defenders need to prepare for a greater volume of more varied attacks *and* risks from employees using new AI technologies, increasing the need for cyber hygiene.

It may soon become possible for attackers to use AI tools to scan environments for exploitable vulnerabilities. This will dramatically improve their reconnaissance capabilities.

Currently, the complex decision-making capabilities of LLM-based autonomous agents are relatively rudimentary, but research in this area is ongoing. It is expected that well-funded adversary groups will build upon this research to augment their capabilities. Jude Sunderbruch, Director of the US Department of Defense Cyber Crime Center, has described this as a looming "battle between AI and counter-AI."[3] Harnessing AI in new and more sophisticated ways will continue to challenge defenders going forwards.

Many security and enterprise leaders are already thinking about how to mitigate these risks.

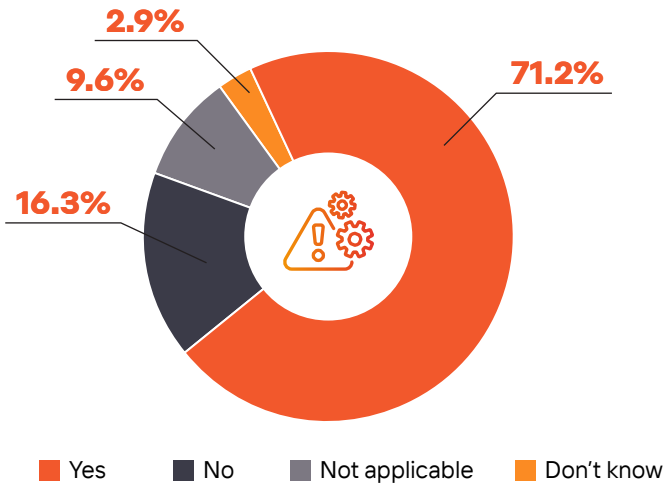
Shadow AI: A Growing Challenge

It takes little effort for employees to adopt publicly-available text-based generative AI systems to increase their productivity. This opens the door to "shadow AI"—use of popular AI tools without organizational approval or oversight. Resulting security risks such as inadvertent exposure of sensitive information or intellectual property are an ever-growing concern.

[3] U.S. Department of Defense, "Battle Looming Between AI and Counter-AI, Says Official," 25 January 2024, Available at: https://www.defense.gov/News/News-Stories/Article/Article/3656926/battle-looming-between-ai-and-counter-ai-says-official/

71% of organizations have already taken strides to reduce risks associated with the adoption of AI.

Has your organization taken any steps specifically to reduce the risks of using AI within its applications and computing environment?



This finding is good news. **Even as enterprises compete to get as much value from AI as they can, as quickly as possible, they're tempering their eager embrace of new tools with sensible caution.**

Still, responses varied across roles. Security analysts, operators, administrators, and incident responders are less likely to have said their organizations had taken AI risk mitigation steps than respondents in other roles. In fact, 79% of executives said steps had been taken, and only 54% of respondents in hands-on roles agreed. It seems that leaders believe their organizations are taking the needed steps, but practitioners are seeing a gap. What's being sold to business decision-makers isn't being seen on the front lines of cyber defense.

Among regions, Asia-Pacific has the highest percentage of respondents (83%) claiming their organizations have taken steps to limit the risks that come with using AI.

Leaders believe their organizations are taking the needed steps to mitigate AI risk, but practitioners are seeing a gap.

Best Practices for Securely Implementing AI in the Enterprise

As organizations embrace generative AI and LLMs, CISOs and security teams must ensure their reliability and mitigate the risks associated with their use. AI security should be baked into the AI adoption journey from the very start.

In the simplest of terms, AI is composed of three elements:

Data, regardless of how or where it is stored

Compute resources

AI algorithms and models, which are sets of mathematical equations, technical computations, and instructions on how to use the data

---

These elements need to be secured like anything else within the digital environment. AI should be thought of as an extension of the attack surface and digital estate.

Security teams should apply the same principles to their data storage and AI models that they would to any technology in their organization. These include:

- Extended visibility
- Continuous monitoring
- Detection and response
- Access controls
- Defense-in-depth
- Threat vulnerability management
- Zero trust
- Control

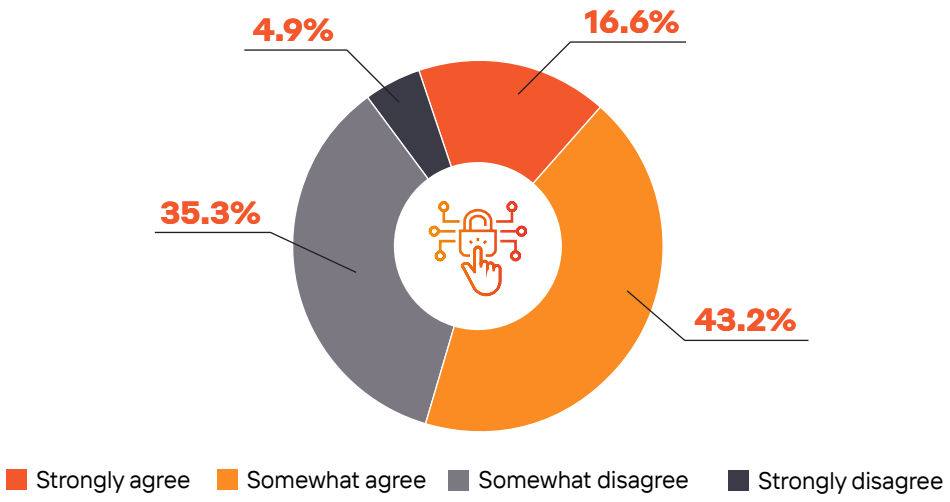
Additional elements are unique to AI security:

- Develop a robust Testing, Evaluation, Verification and Validation (TEVV) plan
- Conduct red-teaming or exhaustive testing of your AI models, data storage, and APIs to identify vulnerabilities and mitigate risk
- Train security teams in Adversarial Machine Learning (AML) techniques and embed these within AI development initiatives

Just five percent of security professionals have strong confidence in their organizational preparedness to face the next generation of threats.

Describe your agreement with this statement:

"Based on the security capabilities currently in place, my organization is not adequately prepared to defend against AI-powered threats and attacks."



Self-doubts are rampant. A majority of respondents (six out of every ten) believe their organizations are inadequately prepared to face the next generation of AI-powered threats.

What may be even more telling is that just five percent have strong confidence in their organization's readiness to defend against AI-powered threats and attacks.

Comparing respondent job titles, we find that security administrators are the most skeptical (with 72% stating their organizations are not adequately prepared), which is consistent with the fact that their hands-on experience gives them a more immediate understanding of the size and scope of the challenge.

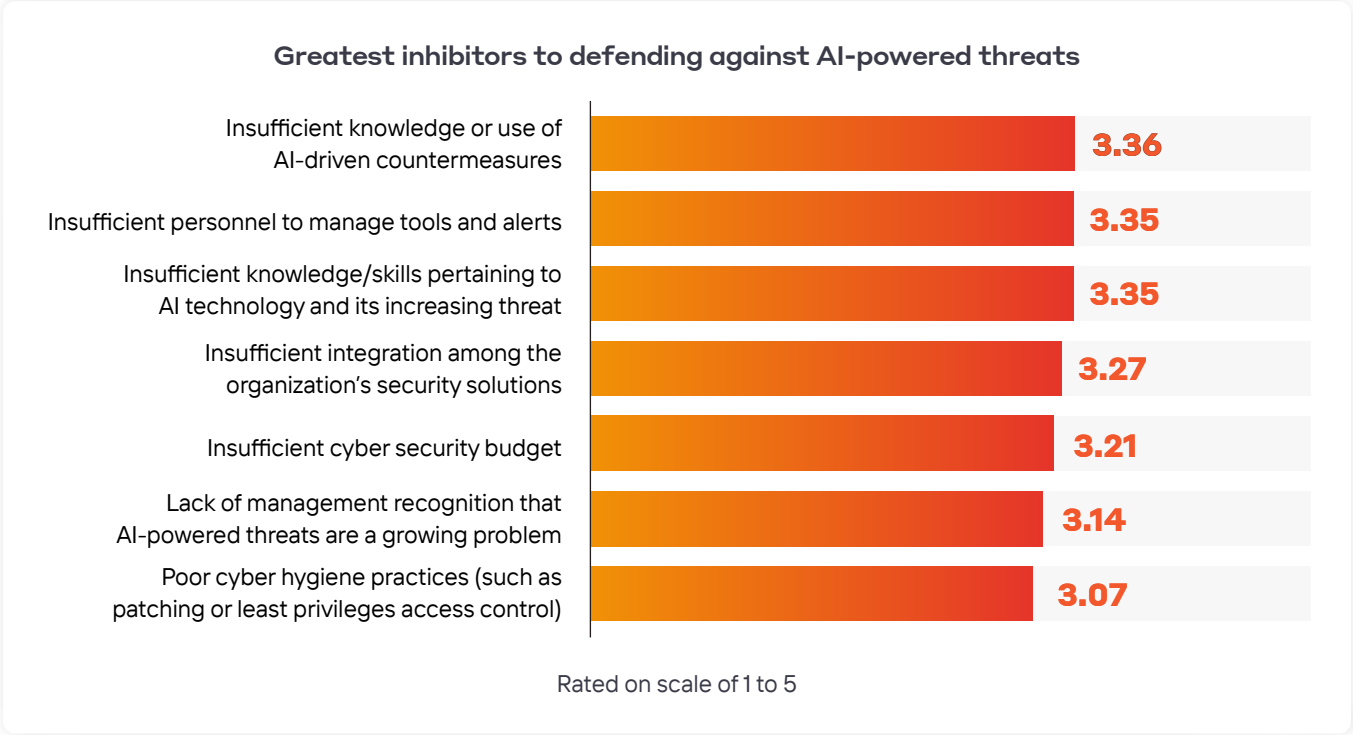
Respondents in mid-sized organizations (5-10K employees) feel they are the least prepared, while those in the very largest companies (more than 25K employees) feel they are the most prepared.

Survey participants in Asia-Pacific are most likely to believe their organizations are unprepared (with 80% agreeing), while those in Latin America feel they're most prepared (46% agreement). This is in keeping with the earlier finding that greater numbers of respondents in APAC perceive AI-powered threats to be imminent.

This finding may reflect a trend that threat researchers have observed elsewhere. In the IBM X-Force Threat Intelligence Index, Asia-Pacific was the most-impacted region in 2021 and 2022, with Europe closely trailing as the second-most-impacted. In 2023, Europe moved into the top spot, accounting for 32% of incidents the researchers observed. That same year, North America represented 26% of incidents and Asia-Pacific 23%, while Latin America experienced only 12%.[4] It's possible that Latin American respondents are currently feeling more optimistic about their preparedness because they're seeing lower threat volumes. This could change suddenly and without warning.

[4] IBM, X-Force Threat Intelligence Index 2024, Available at: https://www.ibm.com/downloads/cas/L0GKXDWJ

Lack of knowledge and personnel are the biggest barriers to defending against AI-powered threats.



Inhibitors to Effective Defense

The top-ranked inhibitors center on knowledge and personnel. In an industry confronting a worldwide shortage of approximately 4 million professionals,[5] it is only logical that organizations struggle to find people to manage their tools and alerts. This challenge will only grow more pressing as the adoption of AI enables attackers to launch more attacks faster.

It is no surprise organizations are having trouble finding professionals with the skills to manage AI-powered defenses. Adding this additional domain has changed the demands on security teams. It's not that AI tools require more people to operate, but rather that practitioners need broad and deep knowledge across a solution stack that is undergoing rapid change. This makes end user education more important than ever.

Finding people who understand how AI is and will be used in security should continue to be a top priority for organizations for some time to come. In particular, professionals who can operationalize the output of machine learning models for efficacy and accuracy will be in very high demand.

Our survey also revealed a lack of understanding of AI-driven threats. Cyber security professionals—and the tools they're operating—aren't keeping pace with the latest changes in attacker behavior and tactics.

Especially noteworthy is the problem posed by a lack of integration among security solutions.

Insufficient budget was near the bottom of the list of inhibitors, suggesting that stakeholders who make funding decisions are taking note of these issues.

Relying on siloed point solutions makes achieving organization-wide visibility near impossible, but it also makes security operations more time- and labor-intensive, exacerbating the skills shortage.

[5] ISC2, Cybersecurity Workforce Study 2023, Available at: https://media.isc2.org/-/media/Project/ISC2/Main/Media/documents/research/ISC2_Cybersecurity_Workforce_Study_2023.pdf?rev=28b46de71ce24e6ab7705f6e3da8637e

POWERING UP

The Impact of AI on Cyber Security Solutions

Many of the biggest operational challenges cyber security leaders currently confront are familiar.

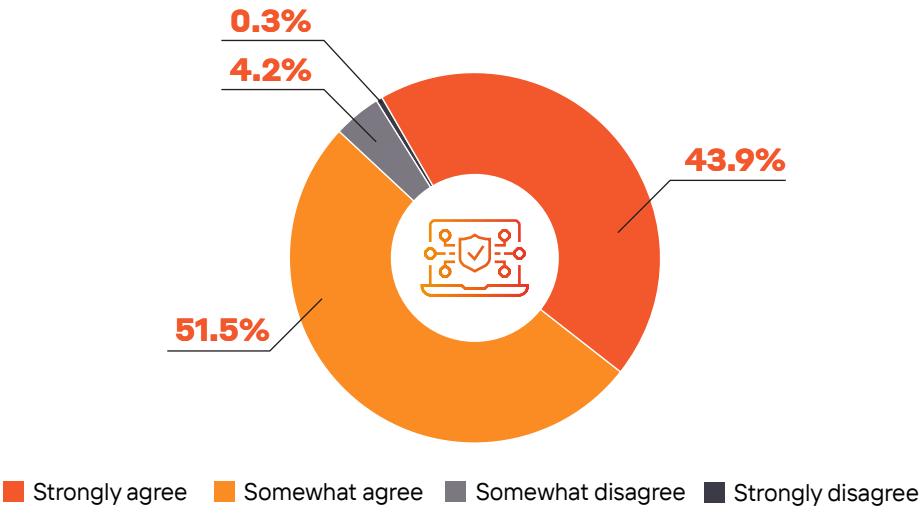
Overwhelming alert volumes, high false positive rates, and endlessly innovative threat actors keep security teams scrambling. Defenders have taken a reactive approach, struggling to keep pace with an ever-evolving threat landscape, and it makes sense that they would do so. It is hard to find time to address long-term objectives or revamp operational processes when you are always engaged in hand-to-hand combat.

The impact of AI on the threat landscape will soon make yesterday's approaches untenable. Cyber security vendors are racing to capitalize on buyer interest in AI by supplying solutions that promise to meet the need. But not all AI is created equal, and not all these solutions live up to the widespread hype.

95% of cyber security professionals agree that AI-powered solutions will level up their organization's defenses.

Describe your agreement with this statement:

"AI-powered security solutions significantly improve the speed and efficiency of my organization's ability to prevent, detect, respond to and recover from cyber-threats."



Stakeholders firmly believe they'll need to adopt AI to keep pace with evolving threats.

Not only is there strong agreement about the ability of AI-powered cyber security solutions to improve the speed and efficiency of prevention, detection, response, and recovery, but that agreement is nearly universal, with more than 95% of respondents affirming the above statement.

This is about much more than generative AI. There are many other types of AI, and many other use cases where they'll prove invaluable to cyber defense.

Agreement was broadly consistent across all demographics, geographies, and organization sizes.

Improving threat detection is the #1 area within cyber security where AI is expected to have an impact.

AI is poised to transform not just the threat landscape but the solution landscape as well, a fact defenders understand. **Threat detection is their biggest need by far. Unfortunately, they're less than clear on exactly how AI can most effectively be applied here or elsewhere.**

The most frequent response to this question, improving threat detection capabilities in general, was top-ranked by slightly more than half (57%) of respondents. This suggests they hope that AI will rapidly analyze enormous numbers of validated threats within huge volumes of fast-flowing events and signals. And that it will ultimately prove a boon to front-line security analysts. They are not wrong.

Identifying exploitable vulnerabilities (mentioned by 50% of respondents) is also important. Strengthening vulnerability management by applying AI to continuously monitor the exposed attack surface for risks and high-impact vulnerabilities can give defenders an edge. If it prevents threats from ever reaching the network, AI will have a major downstream impact on incident prevalence and breach risk.

We will see later in this report that generative AI is often misunderstood by security leaders and practitioners, and actually has little to no role to play in threat detection and proactive attack surface management.

50% of security stakeholders believe that AI will have the biggest impact on vulnerability identification.

The domains where defensive AI is expected to have the greatest impact:

61%
cloud security

50%
data security

46%
network security

The Role of Generative AI in Cyber Defense

Where generative AI *can* help:

- Accelerating the data retrieval process within threat detection (since it makes it possible for analysts to interact with algorithms using easy-to-understand natural language prompts)
- Creating quick incident summaries
- Automating low-level tasks in security operations
- Simulating phishing emails and other attack tactics

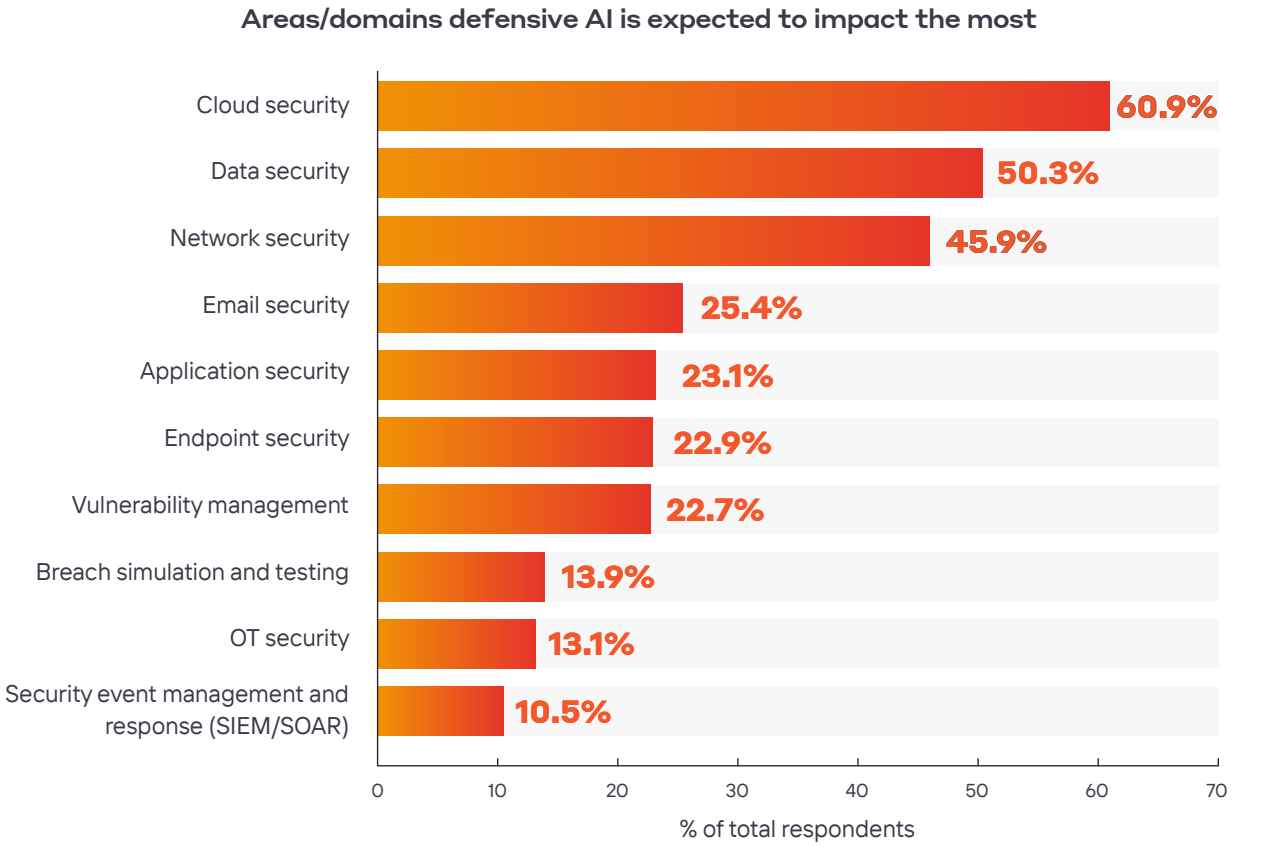
Interestingly, most of these use cases were ranked lower by survey participants.

An Expanding Array of Use Cases

Survey participants may not understand how AI can help detect new unknown or zero-day threats. Because attackers may someday use the technology to create new malware (including custom exploits), it is vital that defensive strategies move away from signature-based approaches. Unsupervised machine learning algorithms can do exactly this, determining which traffic patterns and behaviors are normal for an organization and then immediately flagging anything that deviates from this baseline to highlight misuses, abuses, and misconfigurations. Because this type of AI is constantly revising its assumptions about behavior, it's perfectly suited to address the next generation of shifting threats.

For AI to live up to its potential, however, it must work with precision. If it is prone to delivering false positives, it'll only create more noise, not the risk- and labor-reduction defenders need.

Cloud security (61%), data security (50%), and network security (46%) are the domains where defensive AI is expected to have the greatest impact.



Respondents selected broader domains over specific technologies. In particular, they chose the areas experiencing a renaissance. Cloud is the future for most organizations, and the effects of cloud adoption on data and networks are intertwined. All three domains are increasingly central to business operations, impacting everything everywhere.

Our threat research shows that the most commonly observed patterns in breaches involved multiple steps of lateral movement, often with cloud and SaaS solutions as initial entry points. In the second half of 2023, the second-most-frequently observed Cyber AI Analyst activity pattern was SaaS Hijacking (categorized as Privilege Escalation within the MITRE ATT&CK knowledge base).[6]

As enterprises become more reliant upon SaaS applications and cloud platforms—and attack surfaces expand—this trend will likely persist. Attackers will continue to target cloud environments, seeking unauthorized access to privileged accounts through account takeovers. Once they've hijacked these accounts, they can perform a variety of nefarious activities including data exfiltration or launching phishing campaigns.

Going forward, it will be paramount for organizations to augment their cloud and SaaS security with AI-powered anomaly detection, as threat actors sharpen their focus on these targets.

Responses were remarkably consistent across demographics, geographies, and organization sizes, suggesting that nearly all survey participants are thinking about this similarly—that AI will likely have far-reaching applications across the broadest fields, as well as fewer, more specific applications within narrower categories.

[6] Darktrace, End of Year Threat Report: Analysis of the Second Half of 2023, Available at: https://assets-global.website-files.com/626ff19cdd07d1258d49238d/65c10e74dd798acb4fc99c72_End_of_Year_Threat_Report_2023_Final.pdf

Most security stakeholders (71%) are confident that AI-powered security solutions are better able to block AI-powered threats than traditional tools.

Confidence in Traditional Tools Wavering
Responses to this question belie a degree of uncertainty and ambivalence.

Slightly more than half of the survey participants (51%) are fairly confident that traditional cyber security solutions can detect and block AI-powered attacks. The other half are not very confident in the same thing. And only 15% are extremely confident in the effectiveness of traditional solutions.

To some extent, this reflects the current reality. AI has had a dramatic impact on the email threat landscape, and well-resourced threat groups are now leveraging AI-generated scripts to evade signature-based detection mechanisms. As these techniques grow in popularity, they'll become accessible to less sophisticated attackers, and eventually be standard operating procedure for all cyber criminals.

AI-powered threats are not a simple category. Instead, they span many different attack types—from generative AI being used to create emails and customize malware to more advanced reconnaissance techniques. Each of these different threat "flavors" works in a different way, and each needs to be addressed with a different type of defensive AI.

Mounting Disillusionment
Our survey revealed there is little faith in traditional approaches to detection that rely on analyzing known threats. These approaches cannot stop today's attacks, let alone future ones. The lack of confidence voiced makes clear that a shift is needed, and it's needed now.

66%

of survey respondents agree that AI-powered solutions will be able to stop AI-powered threats automatically.

Implementing AI to Counter AI-Driven Threats

There is strong agreement that AI-powered solutions will be better at stopping AI-powered threats (71% of respondents are confident in this), and there's also agreement (66%) that AI-powered solutions will be able to do so automatically. This implies significant faith in the ability of AI to detect threats both precisely and accurately, and also orchestrate the correct response actions.

There is also a high degree of confidence in the ability of security teams to implement and operate AI-powered solutions, with only 30% of respondents expressing doubt. This bodes well for the acceptance of AI-powered solutions, with stakeholders saying they're prepared for the shift.

On the one hand, it is positive that cyber security stakeholders are beginning to understand the terms of this contest—that is, that only AI can be used to fight AI. On the other hand, there are persistent misunderstandings about what AI is, what it can do, and why choosing the right type of AI is so important. Only when those popular misconceptions have become far less widespread can our industry advance its effectiveness.

In the next section, we'll explore some of these misunderstandings in greater detail.

Building Trust in AI: What's Needed

AI systems are complex by nature. If they are to be useful to humans, they need to produce outputs that are clear and easy to understand. But these outputs must also be accurate and verifiable if teams are to rely upon them.

To be worthy of the trust of security professionals, an AI solution must be:

- **Explainable.** Humans can see exactly how an AI system is "thinking" if it outputs its decision-making roadmap in easy-to-understand natural language.
- **Transparent.** The very existence of the 'black box' model risks eroding trust between humans and AI, and can create compliance concerns. An AI system should show all the questions it asks and conclusions it reaches when reasoning.
- **Controllable**. Security teams should be able to decide on the role humans play in decision-making. They should also be able to customize models and set thresholds to guide *how* decisions are made.

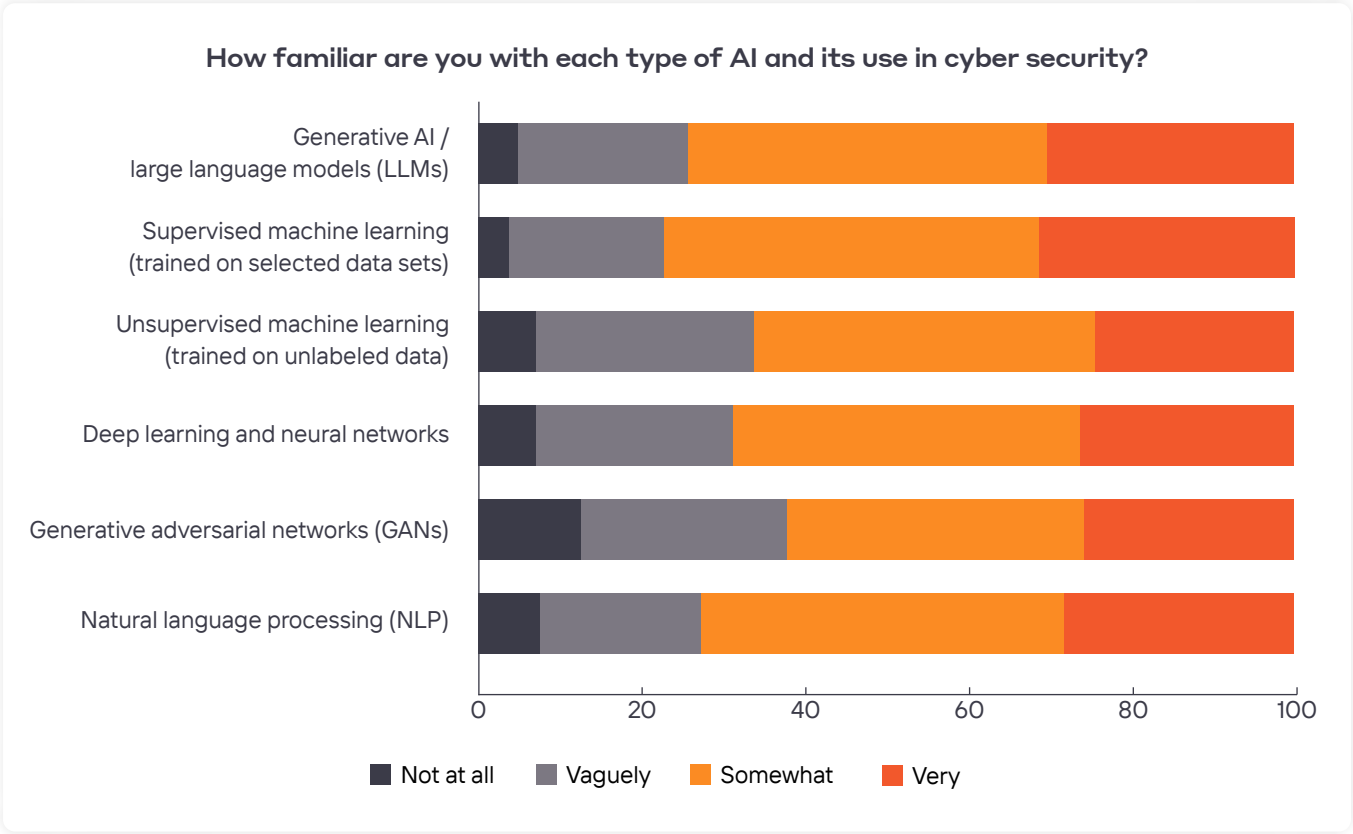
WHAT'S UNDER THE HOOD?

Understanding AI Technologies

A worldwide preoccupation with generative AI may have colored perceptions of what AI is and where it's most effectively applied.

Stakeholders need a more nuanced understanding of the different types of AI if security programs are to maximize the value of AI in real-world operations.

Just 31% of security professionals report that they are "very familiar" with supervised machine learning, the type of AI most often applied in today's cyber security solutions.



It's human nature to believe yourself knowledgeable about technologies that are widely used within your profession. Yet we saw participants admit to remarkably high levels of unfamiliarity with the many types of AI mentioned.

Less than one-third of respondents said they were "very familiar" with the technologies we asked them about. Only 31% reported high levels of familiarity with supervised machine learning (the best-known type of AI). Just 28% said this about natural language processing (NLP) (the second-best-known AI type).

Most survey participants said they were "somewhat" familiar with most of the types of AI technology. Groups of "somewhat familiar" respondents ranged in size from 46% of survey-takers (the largest group, for supervised machine learning) to 36% (the smallest group, for generative adversarial networks (GANs)).

Among the respondents, executives are most likely to view themselves as familiar with AI technologies, as are participants from larger organizations, perhaps because their more specialized roles require greater technical understanding.

Combining the "very" and "somewhat" familiar responses, we see that supervised machine learning (recognized by 77% of survey participants) is the best known, followed by generative AI (recognized by 74% of survey participants), and then NLP (recognized by 73%). With generative AI getting so much media attention, and NLP being the broader area of AI that encompasses generative AI, these results may indicate that stakeholders are understanding the topic on the basis of buzz, not hands-on work with the technologies.

If defenders hope to get ahead of attackers, they will need to go beyond supervised learning algorithms trained on known attack patterns and generative AI. Instead, they'll need to adopt a comprehensive toolkit comprised of multiple, varied AI approaches—including unsupervised algorithms that continuously learn from an organization's specific data rather than relying on big data generalizations.

Types of AI

Different types of AI have different strengths and use cases in cyber security. It's important to choose the right technique for what you're trying to achieve.

Supervised machine learning

Applied more often than any other type of AI in cyber security. Trained on human attack patterns and historical threat intelligence.

Natural language processing (NLP)

Applies computational techniques to process and understand human language.

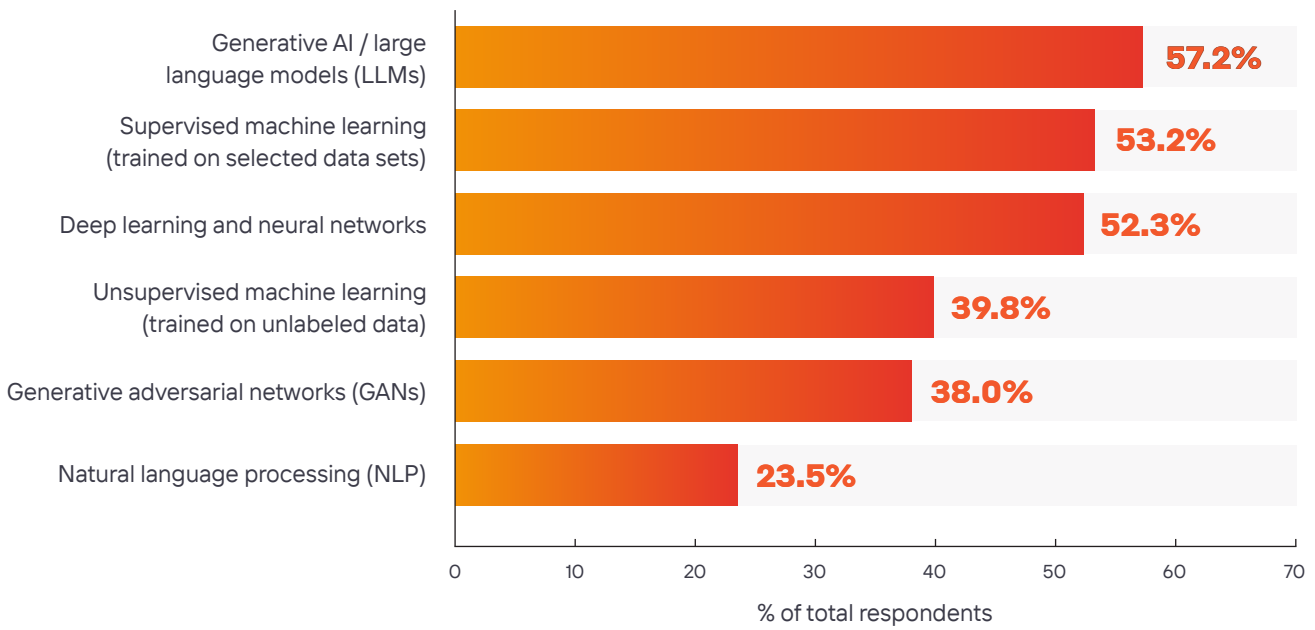
Large language models (LLMs)

Applies deep learning models trained on massively large data sets to understand, summarize, and generate new content. Used in generative AI tools.

Unsupervised machine learning

Continuously learns from raw, unstructured data to identify deviations that represent true anomalies.

More than half of security professionals (57%) believe that generative AI will have a bigger impact on their field over the next few years than other types of AI.



Security stakeholders are highly aware of generative AI and LLMs, so these technologies tend to be top-of-mind when they imagine the future of their field.

But generative AI and LLMs can play only a limited role in cyber defense, though they are very good at certain things. Gen AI is highly capable of:

- Abstracting information from a known environment to augment contextual awareness
- Automating certain tasks, like the creation of incident summaries or sophisticated phishing attacks emulations
- Making it easier for humans to interact with computers

However, LLMs are prone to "hallucinations"—delivering inaccurate or conflicting responses due to errors in training data sets or inconsistencies in the prompts they're given. Plus, prompt-based models are insecure by design. They are vulnerable to prompt injection attacks, which are proving difficult to prevent or defend against. That said, recent strides have been made in securing LLMs against such vulnerabilities as well as improving accuracy.

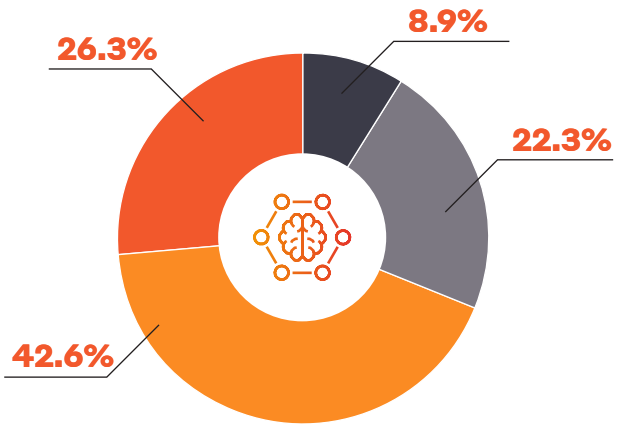
There's no doubt that LLMs have value, but the most effective cyber defenses are those that are built on a combination of different types of AI, with each used in a way that's fit for purpose and that enhances the capability and accuracy of the system as a whole.

Back to School

There's a strong need for AI education across the industry. The more attention generative AI gets, the higher expectations tend to be. As leaders and practitioners discover more about AI, they will need to learn when and where to use it—and how to offset the potential risks that various models and approaches can bring. **As their learning deepens, there will be less emphasis on generative AI and more clarity about how AI can bring value in other areas.**

Only 26% of security professionals report a full understanding of the different types of AI in use within security products.

Which statement best describes your understanding of the types of AI used in your organization's existing cyber security solutions?



- I'm not sure which types of AI are being used; vendor claims are confusing
- I'm not sure which types of AI are being used; I suspect it's mostly generative AI
- I fully understand which types of AI are being used; it's mostly generative AI
- I fully understand which types of AI are being used; it varies by vendor and product

Confusion abounds in today's marketplace. According to our survey, only about one-quarter of respondents (26%) fully understand which types of AI are being used in their security solution stack.

Approximately one-third of respondents (31%) report that they're unsure which types of AI are being used, and/or that they're confused by vendor claims.

Nearly two-thirds (65%) of survey participants believe that what's being used in cyber security solutions is mostly generative AI. In reality, generative AI isn't useful—or used—in threat detection except for identifying phishing emails. These responses likely represent a misperception of how widely types of AI *other than generative AI* are actually being applied in cyber security.

This finding illustrates that while the need for using AI more expansively is enormous, there's too much focus on generative AI. There's also a great gap between what users expect and what vendors can deliver.

Other key results:

- Executives and managers report higher levels of understanding than practitioners.
- Survey participants working for larger organizations report higher levels of understanding than those in smaller organizations. This may be because bigger companies employ more people in differentiated roles, leading to greater specialization and higher levels of knowledge.

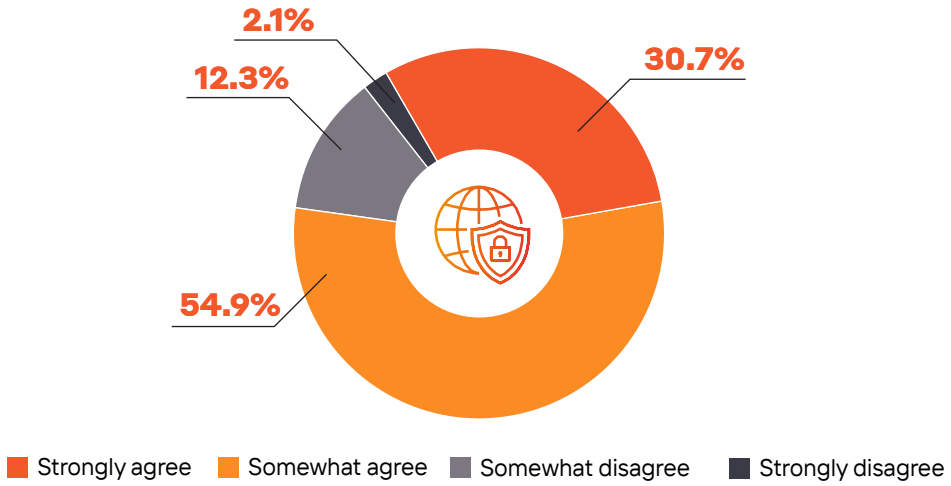
One thing is clear: as the AI revolution unfolds, the speed at which vendors are introducing new AI-powered solutions far outpaces the rate at which practitioners are being trained how to use them.

There's a strong need for greater vendor transparency, as well as more education for end users so they can better understand the technologies being deployed. This will be key for maximizing their real-world value.

Generative AI alone is not enough to stop unknown and zero-day threats.

Describe your agreement with this statement:

"The inclusion of generative AI alone will not significantly improve the ability of a security solution to prevent the impact of new unknown and zero day threats. Other types of AI are also needed to effectively address such threats and attacks."



A significant majority of survey participants (86%) agree that generative AI alone is not enough to address novel threats. This finding was consistent across all geographies, organization sizes, and roles, although executives are slightly less likely to agree than respondents in other roles. Participants in Asia-Pacific were slightly more likely to agree than average, and those in the U.S. slightly less so.

At first blush, this finding seems to contradict our earlier finding: that respondents expected generative AI to impact security more than any other type of AI. However, survey participants likely do understand that generative AI has a limited set of security use cases and is at its most effective when applied in conjunction with other types of AI. This finding reinforces the need for vendor transparency and the need for different approaches to threat detection using AI than what is widely seen today.

It's vital that stakeholders evaluating prospective solutions understand how AI-powered solutions work. This way, they can see where AI can be used most effectively. They should also evaluate whether they're really getting advanced AI-powered detection, or whether solutions are still based on yesterday's methods of understanding existing threats first.

This survey illustrates that stakeholders understand that yesterday's methods are not what's needed.

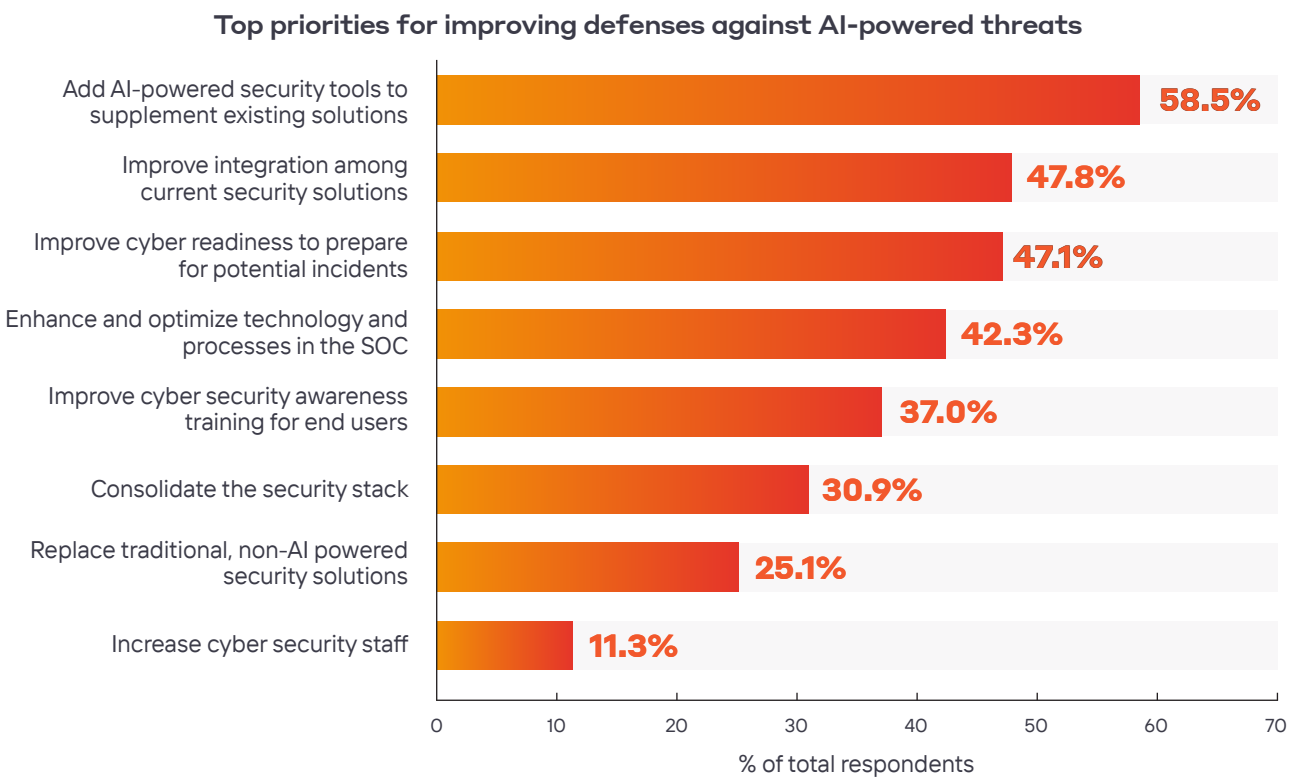
86% of survey participants agree that generative AI alone is not enough to address novel threats.

PREPARING FOR WHAT'S TO COME

Priorities and Objectives

Although security stakeholders are aware that the rise of AI will require them to implement new tools and deploy more advanced capabilities in certain areas, they still entertain multiple different—and sometimes conflicting—opinions about planning for the future.

According to security professionals, their top priorities for improving their ability to defend against AI-driven threats include adding AI-powered tools to their solution stacks and improving toolset integration.



Building on Present-Day Foundations

Generally speaking, organizations are less interested in addressing the challenge of AI-powered threats by increasing staffing or replacing their existing non-AI-powered cyber security solutions.

They're more interested in augmenting their solution stacks with additional AI-powered security tools. This option was chosen by nearly three in five survey participants (59%). Stakeholders are also looking to improve integrations among their existing tools—an option selected by almost half (48%) of respondents.

Many security teams are looking to their existing vendors first when thinking about adding AI-powered tools to their solution stack. This may be because:

- It takes more time and effort to replace existing tooling than it does to add onto the exiting stack.
- Trust has already been established within existing relationships. As long as this is valued, there will always be a need to integrate AI and non-AI solutions.

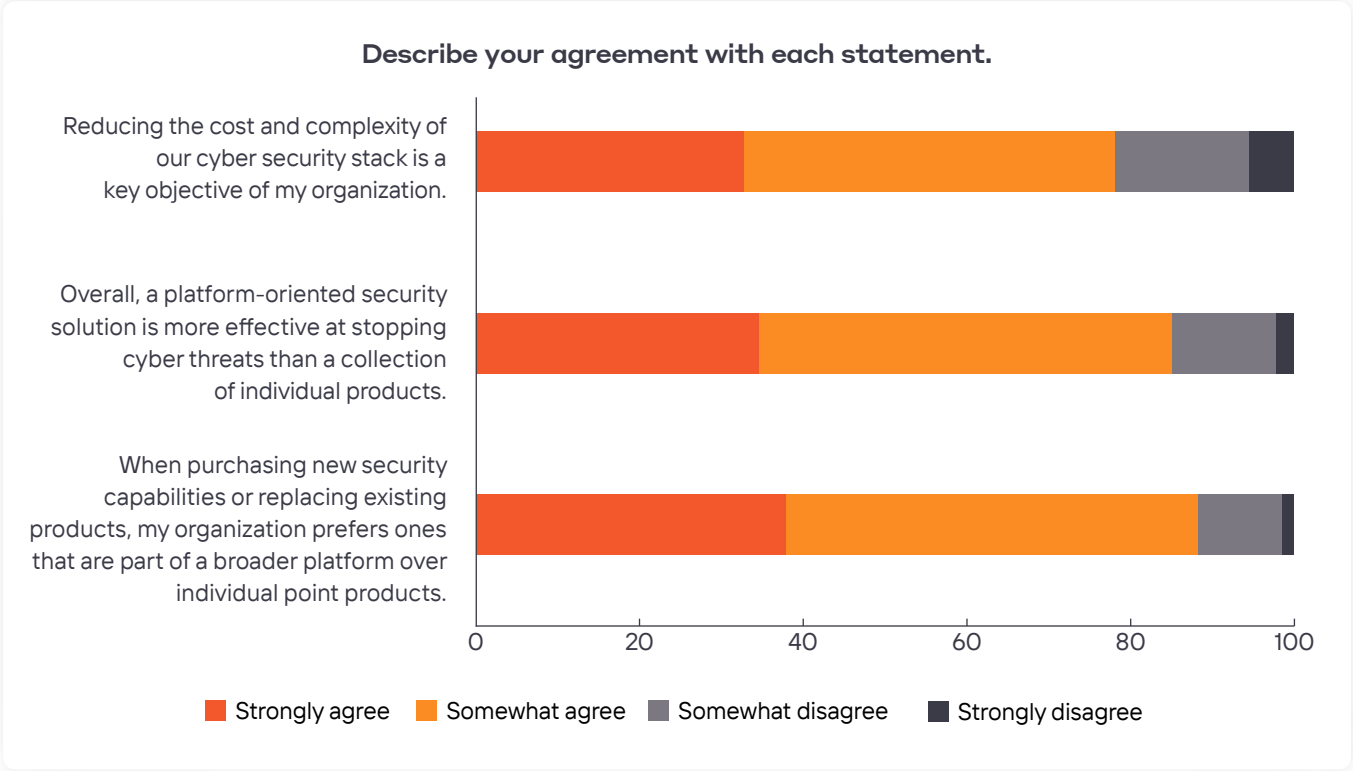
Integrations are in fact key for operationalizing all forms of AI (including generative AI used in incident reporting or alert triage workflows). Integrations are also essential for improving visibility, which is vital as more threats arrive faster.

The key question this raises for security leaders is about solution consolidation. Greater efficiencies can be achieved by adopting a platform approach, especially one that combines multiple types of AI within a single product suite where the technologies can work together.

While each type of AI can work with the others to deliver more value for its respective use case, a synergistic effect can be achieved with product replacement. At what point can AI solutions do everything traditional cyber security tools do, but better?

Eventually, replacement will become inevitable, given the industry's general approach, where new platform solutions tend to achieve greater efficiency than older point products. Exceptions will be few, found only in highly specialized microsegments of the security market.

88% of cyber security professionals prefer a platform approach over individual point products.



Platforms Over Point Products

Respondents expressed a strong preference for a platform-centric approach in their cyber security solution stacks. This is undoubtedly due to a far-reaching desire to reduce cost and complexity. A full 78% of survey participants agree this is a key objective for their organizations.

Respondents also expressed confidence that platform-oriented approaches were more effective at stopping threats than collections of point products, with 85% of survey participants agreeing.

Even more widespread was agreement that organizations prefer to purchase new security capabilities within a broader platform rather than as individual point products. As many as 88% of survey participants expressed this preference.

Executives and managers are even more likely than the average respondent to prefer a platform approach, though the preference was roughly consistent across all participant cohorts.

Respondents from Asia-Pacific were more likely to agree that it's important to simplify the security stack than those from other areas. They were also more likely to believe that platforms outshine point solutions for stopping threats.

78% of survey participants agree that a platform-centric approach in their cyber security solution stacks is a key objective for their organizations.

The Need to Apply the Right AI to the Right Cyber Security Challenges

There's tremendous excitement about the growing potential—including generative AI—to augment the skills and capabilities of security teams. But there's also a great need for stakeholders to better understand these emerging technologies. As threat actors continue to advance their capabilities, it will take multiple types of AI applied in the right use cases to fight back. At the same time, all these types of AI will need smart integration to reduce cost and complexity while enhancing visibility.

Darktrace's approach to cyber defense is built upon the foundational belief that the right type of AI must be applied to the right use cases. AI is central to Darktrace's approach because it plays a pivotal role in identifying novel cyber threats that most other solutions miss. But Darktrace combines multiple AI methods, which work together and reinforce one another. Darktrace builds technology by looking at where AI can best augment the security team and where it can be used responsibly to have the most positive impact on their work.

Darktrace ActiveAI Security Platform

- Business-centric unsupervised machine learning, which can spot and stop entirely new threats as soon as they are launched, even if they've never been seen before. (This includes neural networks, clustering methods, regularization, and probabilistic anomaly detection).
- Bayesian probabilistic methods allow models to be efficiently updated and controlled in real time.
- Applied supervised machine learning is used in Cyber AI Analyst, an investigation tool trained on years of expert analyst behaviors. Cyber AI Analyst was created to greatly expedite triage and reduce time-to-understanding for security teams.
- NLP is used to explain and clarify the decision-making process undertaken by AI agents, providing transparent evidence of how they arrived at their conclusions.
- LLMs categorize malicious communications based on textual properties and build out a heuristic understanding of hostname properties. They're also used in sophisticated attack emulations.
- Deep learning engines replicate the thought processes of humans.
- Graph theory can understand and illuminate the incredibly complex relationships between people, systems, organizations, and supply chains.

In the age of AI, attackers will continue to advance their skills and capabilities to launch a new wave of prolific and increasingly sophisticated threats. As they do so, defenders will no longer be able to rely on yesterday's tools—like supervised machine learning—or on cobbled-together sets of point solutions. Instead, organizations must apply the right types of AI in the right areas within their security stacks to best position themselves for the future.

Through continuous product innovation, Darktrace helps augment human teams to protect against evolving threats in real time—including the latest and most innovative cyber threats. This unique approach facilitates positive and effective human-AI partnerships to meet every organization's unique needs.

Survey Methodology

Our survey was conducted online in January 2024. Respondents held a variety of positions within information security. Nearly one-third (28%) were CIOs, CISOs, or other senior leaders. About one-quarter (26%) were hands-on practitioners such as security analysts, operators, and incident responders, admins, architects, or engineers. Survey participants came from 14 different countries in four different regions, including North America, Latin America, Europe, and Asia-Pacific. Their organizations ranged in size from 500 employees to more than 25,000, with most (57%) working for organizations with more than 1,000 and less than 10,000 employees.

About Darktrace

Darktrace (DARK.L), a global leader in cyber security artificial intelligence, delivers complete AI-powered solutions in its mission to free the world of cyber disruption. Its technology continuously learns and updates its knowledge of 'you' for an organization and applies that understanding to achieve an optimal state of cyber security. Breakthrough innovations from its R&D Centers have resulted in more than 165 patent applications filed. Darktrace employs 2,300 people around the world and protects over 9,200 organizations globally from advanced cyber-threats.

DARKTRACE
Evolving threats call for evolved thinking™

North America: +1 (415) 229 9100
Europe: +44 (0) 1223 394 100
Asia-Pacific: +65 6804 5010
Latin America: +55 11 4949 7696
info@darktrace.com
darktrace.com