December 2025

The State of Human Risk 2025

The New Paradigm of Securing People in the AI Era

The Evolving Workforce Requires Evolved Security

The workforce is transforming at an unprecedented rate and scale. The near-term future is one of humans plus AI agents working in harmony, underpinned by a security program that proactively manages the behavioral risk of both.

We surveyed 4,200 professionals in 14 countries to uncover how they are adapting to this new paradigm. The reality: cybersecurity leaders are facing challenges on multiple fronts. People continue to be targeted by cybercriminals, make mistakes and intentionally exfiltrate data. Email remains the riskiest channel — but other applications, such as Teams and Slack, are catching up.

At the same time, cybersecurity leaders say AI threats are their top challenge when it comes to behavioral risk, and 43% of organizations experienced an increase in AI-related incidents in the last 12 months (the second highest increase, after email).

Organizations need to move fast to secure this new paradigm of people and agents, embracing workforce trust management by fully adopting human risk management (HRM) and extending its principles to secure AI agents.

What's Inside

Human Risk Management Snapshot 2025

96% of organizations struggle to secure the human element

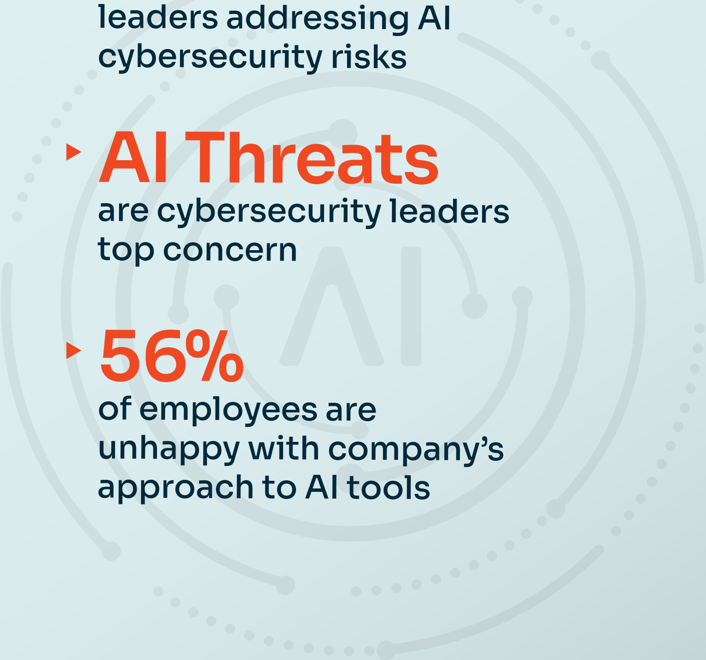
90% increase in incidents relating to the human element

57% increase in email-related incidents

Only **16% of organizations** have a well-established Human Risk Management (HRM) program...

...and only 29% have excellent visibility into human risk

Securing the AI layer is a growing challenge



- **43%** increase in incidents from AI applications

Despite **98%** of cybersecurity leaders addressing AI cybersecurity risks

- **AI Threats** are cybersecurity leaders top concern
- **56%** of employees are unhappy with company's approach to AI tools

Only 6% of employees wouldn't change their organization's cybersecurity program

97% of cybersecurity leaders want more budget to secure the human element

The Human Risk Landscape

Securing people from both external attacks and their own mistakes remains a critical challenge for organizations. In fact, the risk from the human element is growing: 90% of cybersecurity leaders reported an increase in incidents over the last year.

All 700 of the cybersecurity leaders we surveyed had experienced security incidents involving employees in the last 12 months. When categorizing the causes:

- 93% said they occurred due to cybercriminals exploiting employees
- 90% had incidents caused by people making mistakes
- 36% reported incidents because of malicious insiders

A multitude of factors govern people's behavior as they make decisions and interact with applications, systems and data at work — including ownership and accountability.

Most employees sign agreements confirming that the information they work with, including intellectual property, customer data and business plans, belongs to their organization. However, when we surveyed 3,500 employees, only 53% said their company actually owns that information.

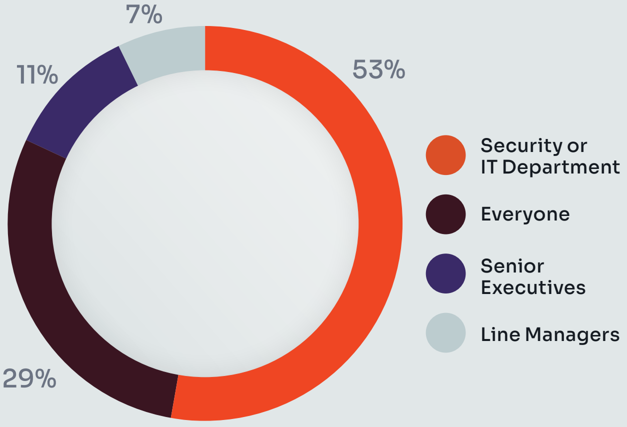
That leaves 47% who believe the data they handle every day belongs to them, their team or their department rather than the business.

For cybersecurity leaders, this creates a mixed picture. A sense of ownership can encourage care and responsibility — but it can also introduce risk if employees start making their own rules about how data is used, stored or shared.

Even more concerning, many employees do not see security as part of their own role. Only 29% believe that everyone is personally responsible for protecting company data. The rest believe it is mainly the responsibility of IT and Security teams (53%), senior executives (11%) or their direct managers (7%).

This gap between security awareness and action leaves organizations exposed. People handle sensitive information every day, but too often they do not realize how much responsibility that carries.

Employees Share Who Has the Greatest Responsibility for Cybersecurity and Data Protection



Letting Outsiders In

As the greatest risk to organizations, we'll examine the ways people are targeted by cybercriminals more closely in Sections 2 and 3. For now, the headlines.

Email phishing remains the primary way that cybercriminals exploit employees, with 64% of cybersecurity leaders saying they've had incidents caused this way. The problem isn't going away: 57% said the number of incidents had increased in the last 12 months.

External attacks led to account takeover (ATO) in 83% of organizations, with more than half of cybersecurity leaders (59%) saying that a phishing email opened the door to cybercriminals.

As well as stealing credentials (52%), cybersecurity leaders said that external attacks aimed to steal data (65%) and financial resources (46%), compromise the supply chain (26%) and provide intelligence for espionage (20%).

Everybody Makes Mistakes

Email is also the primary risk channel when it comes to employees making genuine mistakes (without involvement by cybercriminals). Half (49%) of cybersecurity leaders said they experienced incidents caused by misdirected email. Improper storage (43%) and employees oversharing (e.g. via Teams or Slack) (38%) are the next two biggest causes.

The explosion of AI applications has created a new and complex vector for organizations to secure. These applications experienced the second largest growth in the number of incidents (43%) due to, for example, employees uploading company documents for analysis.

Deliberate Insiders

Just over one-third (36%) of cybersecurity leaders acknowledge that employees have intentionally caused security incidents in the last 12 months. Only 6% of these were successfully stopped before the employee achieved their goal.

Leaking or selling data to a competitor was the top outcome from incidents (43%), followed by leaking data online (37%) and taking data to a new job (35%).

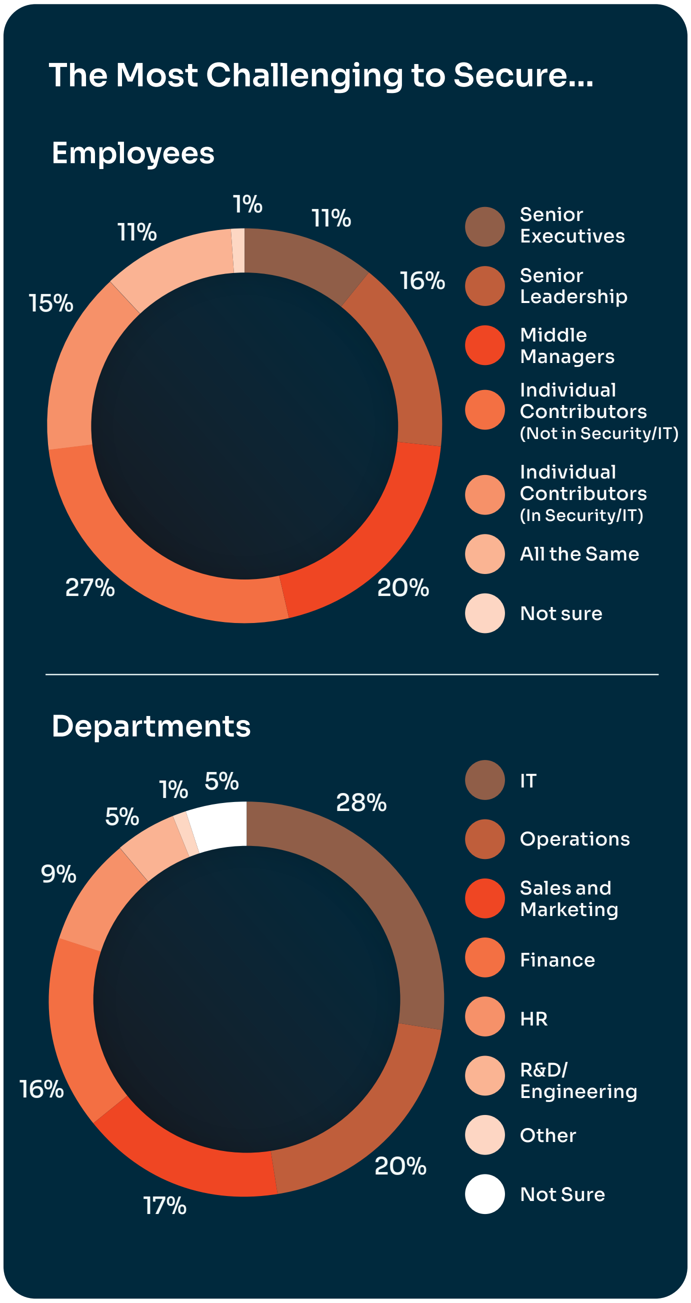
Who Are the Riskiest Employees?

Ninety-six percent of cybersecurity leaders say they find it challenging to secure at least one group of employees in their organization.

Individual contributors outside of the IT or Security teams typically make up the largest part of a workforce, so from a volume perspective, it makes sense that this group comes out as the most challenging to secure (27%).

Together, senior executives and senior leadership are as challenging to secure, which can reflect the need to protect VIPs' often expansive privileged access from threats such as business email compromise (BEC) and whaling.

Interestingly, when asked which department is the most difficult to secure as a group, cybersecurity leaders chose IT. Again, this likely reflects the privileged access for this employee group; rather than a numbers game, here cybersecurity leaders are challenged by a concentration of risk, often in a relatively small-sized team.


The Most Challenging to Secure...

Employees

Departments

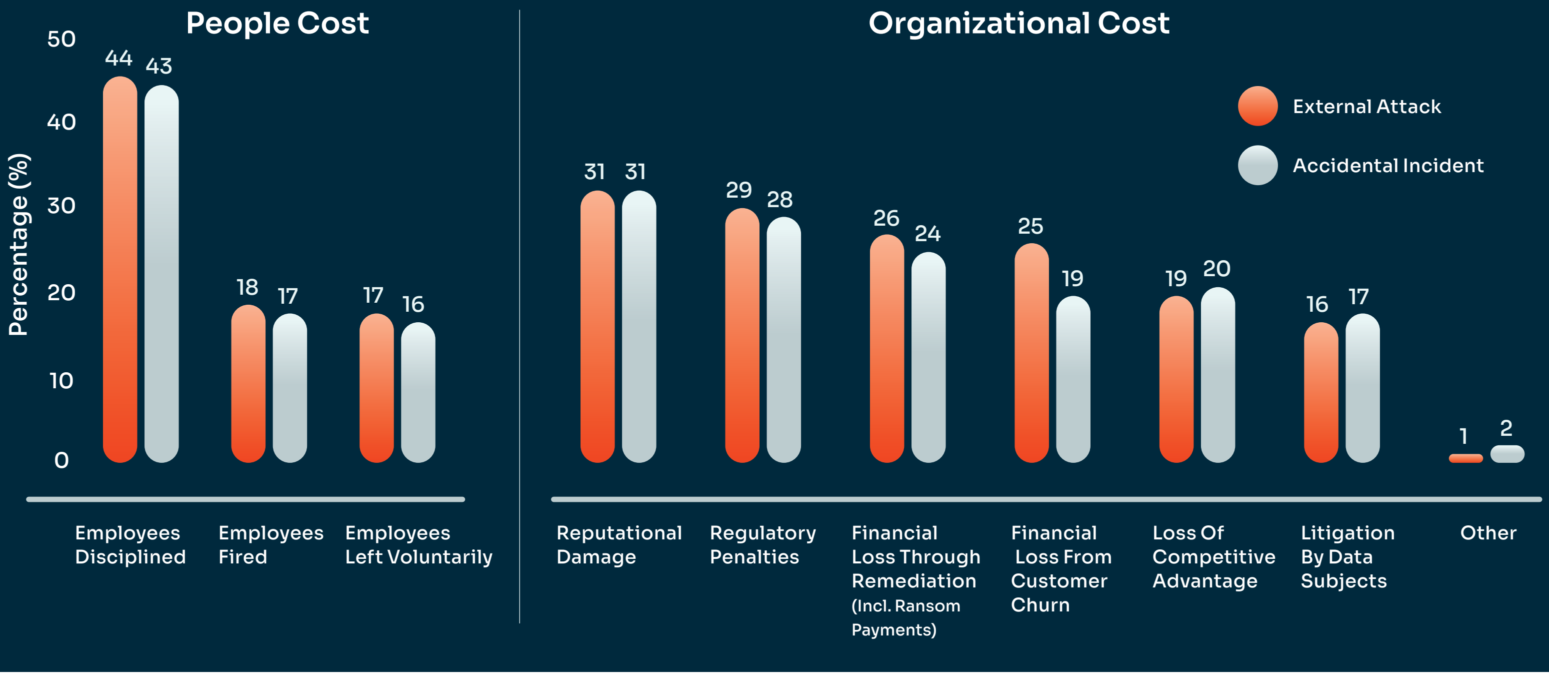
A growing risk: 90% of cybersecurity leaders reported an increase in incidents over the last 12 months

Who Pays for an Incident?

Falling victim to a cybercriminal and making mistakes are inadvertent actions. In both categories, however, people are more likely to be disciplined above any other consequence.

In fact, the outcomes for both types of incidents are remarkably similar, with the top three impacts: employees disciplined, reputational damage and regulatory penalties.

Cybersecurity Leaders Share Outcomes from External Attacks and Accidental Incidents



Outcomes also vary significantly by country, particularly when it comes to disciplining employees. In South Africa, 70% of cybersecurity leaders say they have disciplined employees for falling victim to phishing attacks. Also well above the global average of 44% are Mexico (60%), USA (54%) and Japan (51%).

The results are slightly less stark for accidental incidents. Again, South Africa leads at 56% against the global average of 43%, followed by Mexico (55%), India (51%) and Brazil (50%).

Countries that are least likely to discipline employees are: Denmark (24% for both phishing attacks and accidental incidents), Sweden (30% phishing; 27% accidental), the UK and Ireland (30% phishing; 42% accidental) and France (33% phishing; 34% accidental).

Despite this, cybersecurity leaders in Denmark are slightly more likely to fire employees following a phishing attack (24% vs. 18% global average). Conversely, South African cybersecurity leaders discipline employees more frequently but are less likely to fire them (13% vs. 18%).

Data subjects in the UK and Ireland and Japan are less litigious following an external attack (6% vs. 16% global average), while cybersecurity leaders in Germany, Austria and Switzerland report higher than average reputational damage (43% vs. 31%) and regulatory penalties (40% vs. 29%). Cybersecurity leaders in the USA also experienced more regulatory penalties than average, at 37%.

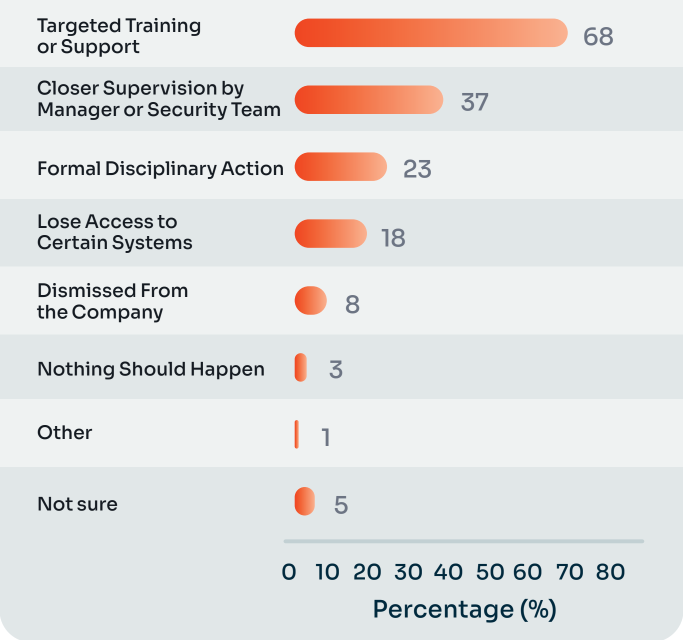
Globally, 39% of cybersecurity leaders say they've fired an employee following an intentional incident. The next two most common outcomes are reputational damage (34%) and employees being disciplined (34%).

A Punitive Disconnect: Employees Favor a Softer Approach

Sixty-eight percent of employees believe that people who inadvertently cause an incident (including clicking on a phishing link) should receive more targeted support. This is by far the most popular outcome. Next is "closer supervision" (37%) and then formal disciplinary action — which at 23% is less than half versus cybersecurity leaders who think the same (59%).

Only 8% of employees believe that someone should be dismissed from the company under these circumstances.

Employees Share What They Think Should Happen Following an Inadvertent Cybersecurity Incident



Securing the Human Element Remains a Complex Problem To Solve

There's significant complexity in managing an increasing human risk. Cybersecurity leaders need to simultaneously ensure established channels like email remain secure while also adapting to focus on emerging vectors, such as AI and messaging applications. As a result, 97% of cybersecurity leaders feel they need more budget to secure the human element.

It's not simply a question of traditional technical defenses; cybersecurity leaders also need to systemically influence behavior.

The data reveals a perception versus policy paradox, with employees' attitudes towards ownership and responsibility deeply impacting behavior. This is compounded by a difference in opinion over consequences for the most frequent and inadvertent incidents (external attacks and mistakes), with employees favoring greater leniency while organizations appear to treat non-malicious insiders nearly the same as intentional offenders.

Consequently, only 6% of employees say they wouldn't change anything about their organization's cybersecurity program.

Human Risk Management (HRM) is a completely new approach to securing the human element, leveraging data to create individual risk scores for each employee that inform AI-driven defenses to provide personalized security and in-the-moment coaching, which then feeds into organizational policy and processes. In this way, employees become actively engaged in cybersecurity rather than passive "passengers."

As we'll see in Section 5, investment in HRM programs is growing. However, while 16% of cybersecurity leaders would describe their program as "well-established" and 33% say they're currently transitioning to a HRM program, 50% are being left behind.

68% of employees
believe people should be
offered more targeted
training or support
following an inadvertent
cybersecurity incident

*ishing Risk

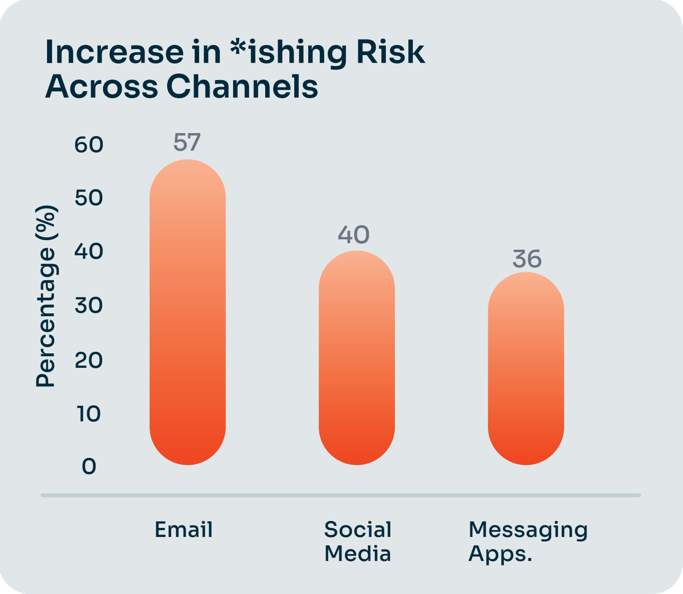
There are many avenues for cybercriminals to target employees — and cybersecurity leaders are feeling this! Ninety-three percent say their organizations had incidents caused by external attacks in the last 12 months.

Email is the riskiest channel, with two-thirds (64%) of cybersecurity leaders reporting incidents occurring this way and 57% saying they've experienced an increase in incidents in the last 12 months.

Messaging applications are the next, with 39% of cybersecurity leaders saying that cybercriminals used applications like Teams and Slack in successful attacks. Messages on social media platforms accessed on corporate devices pose a similar level of risk, at 36%, as did smishing (SMS phishing) (31%).

This signals a shift to boundaryless phishing: increasingly, the threat is everywhere and perimeter-focused defenses are becoming obsolete.

What's more, these "*ishing" risks have increased across every channel.



Threat Alert: Vishing

Vishing remains at relatively low levels versus other payloads but with the rise of deepfakes, it's emerging as part of hyper-targeted kill chains that make the risk both personal and technical.

Our Phishing Threat Trends Report (Vol. 6) revealed 5.3% of phishing emails sent between January 1 - August 31, 2025, contained a phone number as the payload. While this is a relatively small percentage, this was at 0.9% in 2024 — signaling a 449% increase in 2025.

To date, vishing has been tricky for cybercriminals to get right, as it's been difficult to convincingly impersonate someone the target will recognize. All bets are off, however, with the rise of deepfake audio attacks that can turn a cybercriminal's voice into a CEO's.

Vishing has also been reported as a successful tactic in the kill chain for the Scattered Spider criminal gang and their affiliates during their onslaught of attacks on retail and manufacturing giants worldwide.

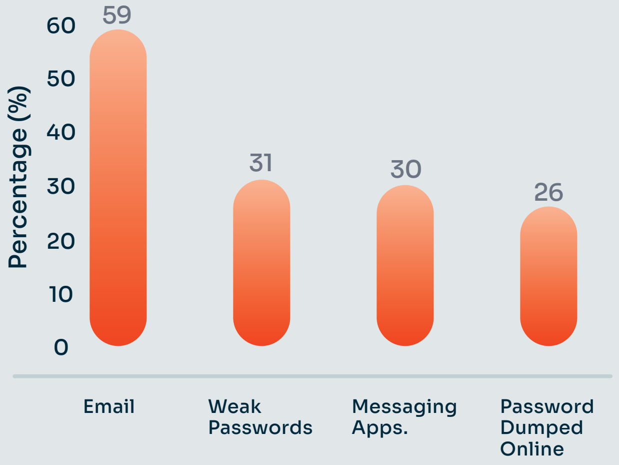
93% of organizations had incidents caused by external attacks

64% of external attacks happened via email

Ask and You Shall Receive

“Asking” employees for their credentials still proves more effective than cracking passwords or finding them online from other breaches. This approach isn’t likely to change any time soon (if ever), so it continues to leave employees susceptible to opening the door to cybercriminals. As attack volume in other applications increases, they’ll compete with email as the primary vector for credential harvesting.

How ATO Attacks Happened Across 83% Of Organizations



Strategic Shifts in Exploitation

It’s our prediction that email will remain the most at-risk channel for at least several more years. However, the rise of multi-channel attacks across messaging applications and vishing, plus a growing trend of cybercriminals exploiting AI tools, means that organizations must also adapt or leave themselves exposed.

An effective HRM program doesn’t just block an attack, it leverages deep behavioral science and threat intelligence to elevate indicators of risk (IoR) to provide early warning signals that shift security into a proactive stance and ensures employees are prepared for threats before they materialize.

Channeling Their Resources

Different channels offer different opportunities to cybercriminals. Email remains key: thanks to its established use, it provides a route to every employee. However, while it’s possible to engineer phishing attacks to bypass certain technologies, such as secure email gateways (SEGs), most organizations have some level of email security in place.

Messaging applications and phones are a little different. It requires more research to determine which platform a company uses for messaging and there’s no way to bulk-send messages to individuals (and groups can arouse suspicion). Similarly, it can be difficult to determine whether a phone number is used for business purposes.

However, these channels are more difficult to secure and, as a result, can be highly successful when used in a targeted attack.

Email opened the door for 59% of account takeover attacks

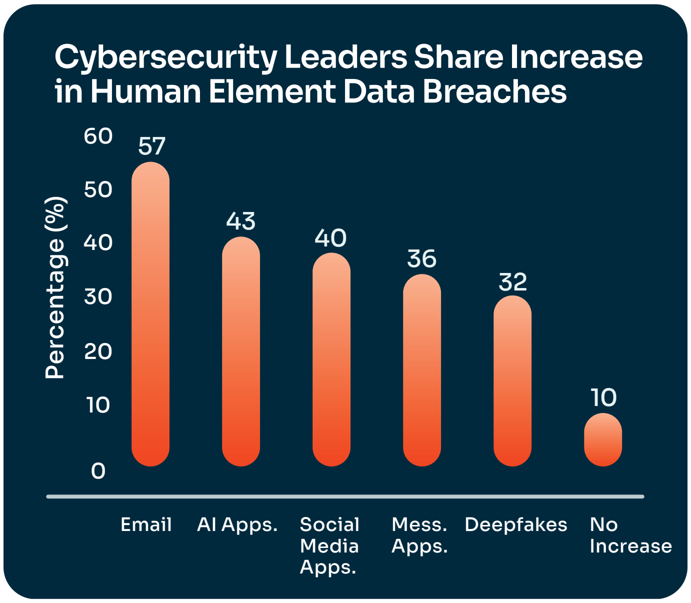
AI: An Uncontrolled Threat Vector

AI is transforming the workforce — and has fast become a critical threat vector.

The rise of autonomous AI platforms and agents introduces new risks, such as undetected manipulation by cybercriminals targeting AI agents, employees oversharing sensitive data and “hallucination” by the applications themselves. Additionally, cybercriminals leverage their own set of AI tools to create more sophisticated attacks at scale.

Cybersecurity leaders are feeling this pain. AI applications experienced the second largest increase in security incidents in the last 12 months, at 43%. Thirty-two percent also say they’ve experienced an increase in incidents related to deepfakes.

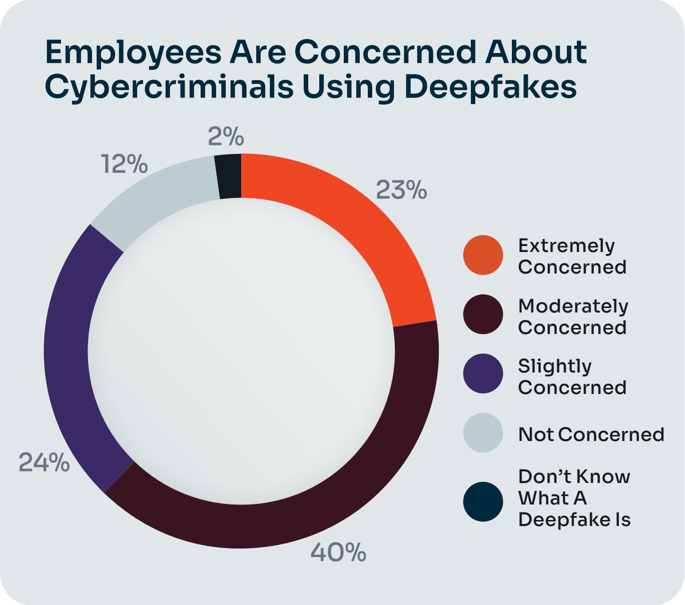
Finally, cybersecurity leaders rank keeping up with constantly evolving AI-powered threats as their greatest challenge when addressing behavioral risk (45%).



AI threats are also on employees’ radars. Eighty-six percent express some level of concern about being tricked by cybercriminals using deepfakes to exploit them into providing access to their company’s data or systems.

This is a brilliant example of internal alignment with the Security agenda that organizations need to capitalize on. Where employees acknowledge they’re concerned they’re also often ready to be protected.

Seventeen percent, however, also say they’ve used an AI tool for work without permission from their Security or IT team.



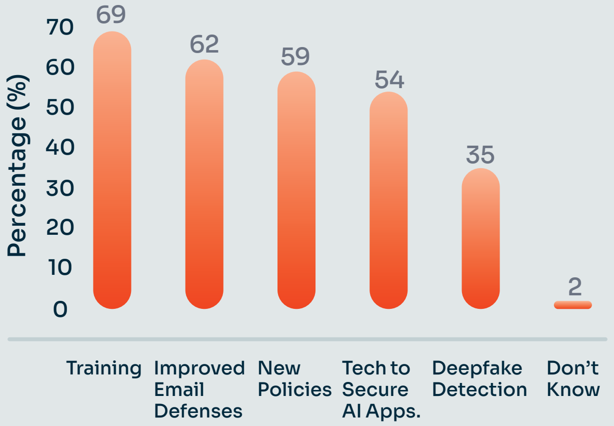
Cybersecurity leaders rank AI-powered threats as their top security risk

86% of employees express concern about deepfakes

What Are Cybersecurity Leaders Doing About All This AI Risk?

Ninety-eight percent of cybersecurity leaders have taken steps in the last 12 months to specifically address AI-related cybersecurity concerns. Training tops the list at 69%, followed by upgrading email security defenses to detect AI-generated phishing attacks (62%). Only one-third (35%) have implemented deepfake detection technology.

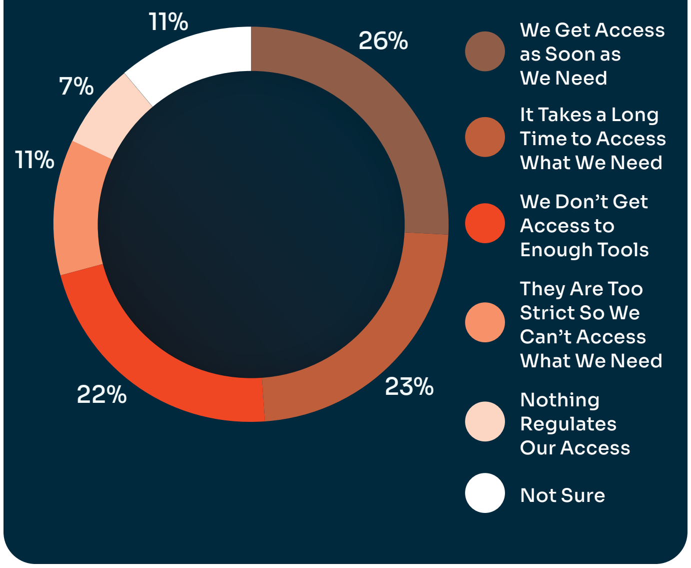
Cybersecurity Leaders Highlight Steps to Curb AI Risk



These measures to simultaneously secure people's access to AI and protect them from AI-powered threats don't seem to be appreciated by employees. Only 26% say their organization gives them fast enough access to the AI tools they need, while 7% worryingly say nothing regulates their access. Over half (56%) feel it takes too long to access the tools they need or they never get access at all.

98% of cybersecurity leaders have taken steps to address AI-related cybersecurity risks

Employees Reveal How They Feel About Their Organization's Approach to Regulating Their Access To AI



The Next Generation of Security: Human + AI Agents

The rapid rise of AI application-based incidents (43% increase) confirms that AI is not just an advanced threat tool for cybercriminals but an immediate, uncontrolled vector for internal employee risk.

Critically, there is an AI governance gap: while 98% of organizations have taken measures to address AI risk, over half (56%) of employees feel access is overly restrictive or slow. This disconnect creates a fertile environment for "shadow AI", which can lead to oversharing by employees and undetected exploitation by cybercriminals targeting AI agents.

The future workforce will encompass both people and AI agents working in harmony — and HRM must evolve to cover the AI layer.

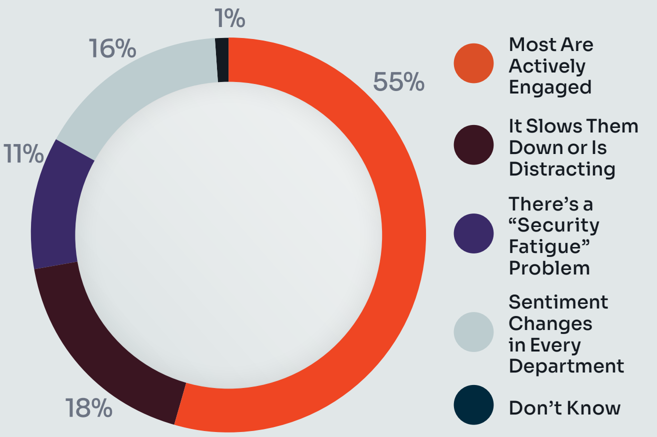
Managing AI risk is a new motion for organizations — but one that must be carried out at pace to avoid critical business activity onto unmonitored, high-risk platforms that ultimately leave organizations exposed.

Is There Still a Culture of Fear?

Culture is central to HRM, which focuses on deeply understanding an individual's risk profile to proactively deliver timely and appropriate security interventions and hyper-personalized coaching before incidents occur. In doing so, employees become active participants in organizational security, rather than "passengers" to corporate policies.

However, only just over half (55%) of cybersecurity leaders consider most employees to be actively engaged in security, leaving 44% who believe it slows employees down or is distracting, they have a security fatigue problem, and engagement levels vary across the organization. Here, an HRM platform would act as a "technological bridge" to transform these employees into active participants.

Cybersecurity Leaders Share Employees' Sentiment Towards Cybersecurity



Employees Mandate for Proactive Security

Ninety-four percent of employees would make changes to their organization's cybersecurity program. The top three are:

- Proactive security tools that stop mistakes before they happen (39%)
- Make training more relevant to each individual's work (39%)
- Greater investment in technology and systems to remove the burden from the employee (33%)

All indicate that employees are fundamentally seeking HRM programs, which use technology to proactively intervene at the point of risk and leverage personalized coaching to increase awareness.

Does the Punishment Fit the Crime?

As we explored in Section 1, the most frequent outcome is for those involved to be disciplined or dismissed.

Globally, 84% percent of cybersecurity leaders believe formal disciplinary procedures are effective, with the majority (59%) saying they are very or extremely effective.

Employees paint a different picture — and one that can erode trust over time if their expectations continue to be misaligned with cybersecurity leaders'.

Only 23% agree people should face disciplinary action for accidental incidents (including phishing) and only 8% say they should be fired.

Instead, employees favor greater leniency and a more people-centric approach to security that focuses on enabling them to work more securely by proactively detecting incidents and providing personalized coaching.

Employees Are Ready for HRM

Critically, employees are not resistant to security but to outdated approaches. Their top requested changes — proactive risk prevention tools (39%) and personalized training (39%) — are direct calls for a modern HRM framework.

To truly activate employees as partners, leaders must shift from a compliance-centric, "punish-and-patch" model to a proactive, technology-enabled HRM strategy that delivers security as an intuitive and relevant work enabler.

Bridging the HRM Execution Gap

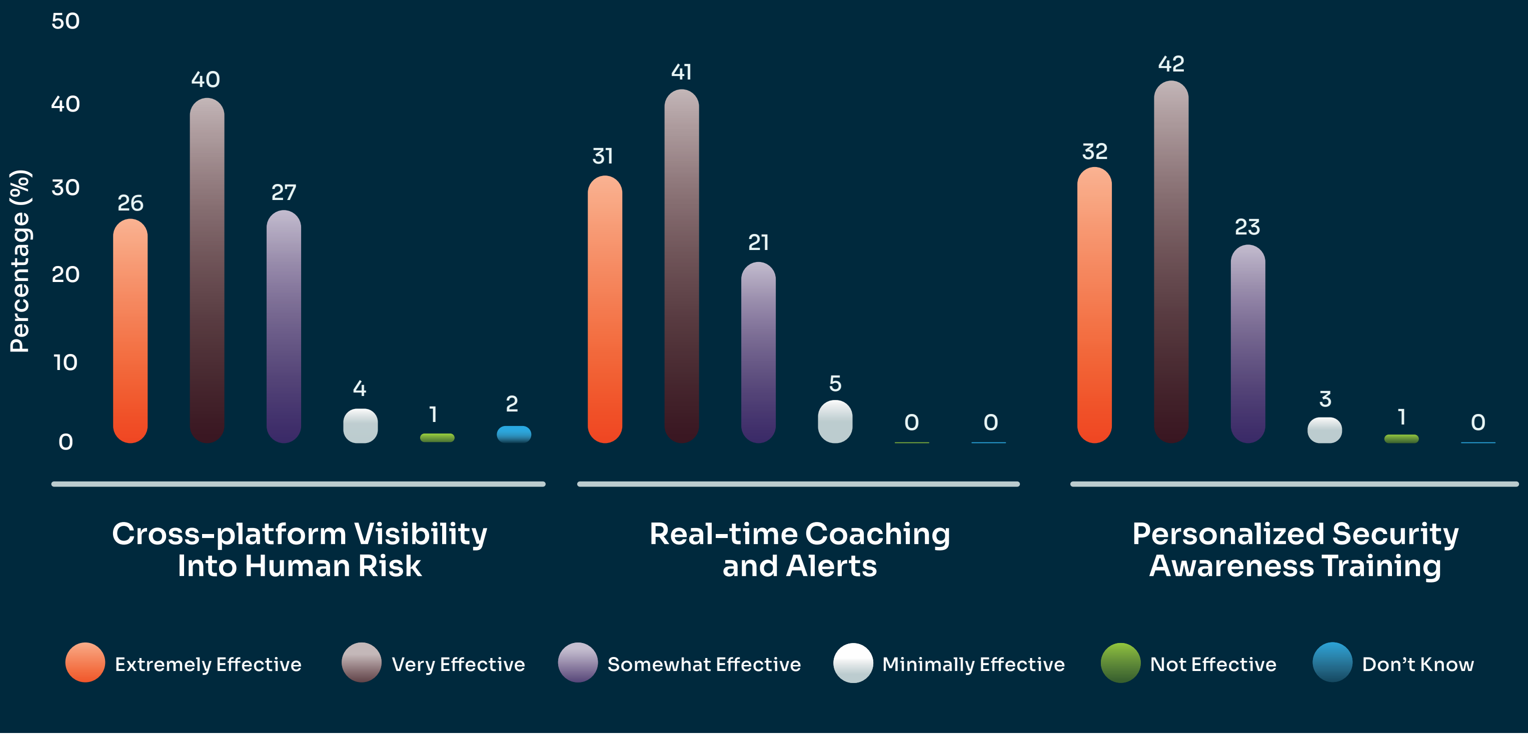
Globally, only 16% of cybersecurity leaders say their organization has a well-established HRM program. This leaves 84% of organizations behind them on the adoption curve, although 33% of these are working to implement one.

Brazil and the USA have the highest adoption rates at 32% each, with a further 44% in Brazil working to adopt HRM and 12% in the USA (with the latter leaving 56% behind).

Adoption is lowest in Japan, Argentina and the UK and Ireland, with 8% per country saying they have a well-established program. In Japan, however, 40% of organizations are working to implement a HRM program, with only 24% in the UK and Ireland saying the same. This leaves 68% of UK and Irish and 58% of Argentinian organizations without even a fledgling HRM program.

Despite this, however, cybersecurity leaders almost unanimously agree on the effectiveness of HRM principles to stop human element security incidents. Ninety-three percent believe that cross-platform visibility into human risk improves HRM, with 66% saying this is either extremely or very effective. Ninety-five percent agree on the effectiveness of real-time coaching and alerts, while 97% say personalized training is effective to some degree.

Cybersecurity Leaders Rate the Effectiveness of HRM Principles



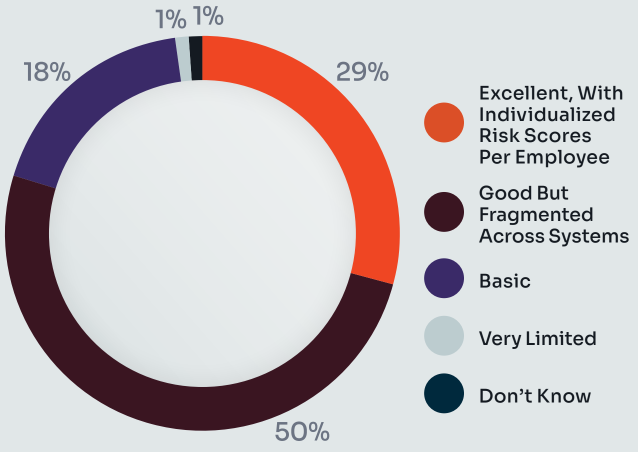
Ninety-six percent of cybersecurity leaders acknowledge they have challenges when securing the human element, with one-third (36%) blaming a reactive approach where risks are predominantly addressed after an incident occurs.

Measuring Risk and Detecting Incidents

Despite 93% of cybersecurity leaders believing that cross-platform visibility improves HRM, less than one-third (29%) say they actually have excellent visibility into human risk, with individualized risk scores per employee, leaving 71% who struggle to know how each person contributes to risk in their organization and tailor their approach accordingly.

Half of cybersecurity leaders (50%) rate their visibility as "good" but fragmented across different systems. About one-fifth (18%) have a very basic level of visibility.

Cybersecurity Leaders Rate Their Level Of Visibility Into Human Risk



At 50% saying they have individualized risk scores, cybersecurity leaders in India are ahead of the curve, followed by the USA (46%) and Denmark (40%). Mexico has the lowest at 16%, followed by the UK and Ireland (18%), New Zealand (20%) and Japan (20%). Japan also has the highest level of basic visibility (e.g. measuring phishing simulation scores in isolation of wider HRM data) at 42%.

With the majority operating with a patchwork view of human risk, it's little surprise that cybersecurity leaders are as reliant on employees informing them about incidents as they are on system detection and alerts.

In external attacks, 56% of cybersecurity leaders say they were made aware of the incident thanks to system reporting and 53% say the person involved told them (some were informed by both). For accidental incidents, the numbers shift slightly to 54% system alerts and 50% by the person involved.

However, can cybersecurity leaders continue to rely on self-reporting if the person knows it may result in formal disciplinary action or even the loss of their job?

For deliberate incidents (which includes malicious behavior and risk-taking despite knowing it breaks policy), the gap is naturally wider. Fifty-four percent of CISOs were made aware by a system alert, while 37% were told by someone who wasn't involved in the incident.

The Battle for Budget

Ninety-seven percent of cybersecurity leaders say they need more budget to effectively minimize employee risk. The top three asks are:

- Monitoring tools for real-time alerts and coaching on risky behaviors (20%)
- Securing AI applications (17%)
- Email security (16%)

The top request recognizes the need for strategic investment in HRM to deliver proactive intervention at the point of risk. Next on the agenda: securing the two channels with the great increase in incidents in the last 12 months. While many organizations have email security budgets in place, securing AI is a new carve-out that must be resourced.

Only 16% of organizations have a well-established HRM program

HRM: Theory vs. Practice

While cybersecurity leaders overwhelmingly endorse the core principles of HRM — specifically cross-platform visibility into risky, real-time coaching and alerts for risky behavior, and personalized training — only 16% have a well-established program.
This disparity creates a strategic vacuum: leaders understand what is effective, but 84% of organizations are behind the curve for implementation.

The lack of actionable visibility is a major obstacle, with most (71%) organizations missing individualized risk scores for each employee. These scores should become part of the fabric of organizational decision making — not just another metric for the security operations center (SOC).

Investing in this approach will bridge the gap between knowing HRM works in principle and proactively stopping incidents.

Less than one-third (29%) of cybersecurity leaders have excellent visibility into human risk

96% of cybersecurity leaders have challenges securing the human element

Securing the Workforce of Today and Tomorrow

Securing the workforce is only becoming more challenging.

Incidents involving people are on the rise and the threat surface is getting wider. Not only do security teams have to contend with protecting established channels, but risk is growing across all communication platforms.

At the same time, AI is rapidly transforming the workforce from one of people to one of people plus agents. Consequently, AI has rocketed up the risk agenda, experiencing the second largest increase in incidents, only behind the well-known threat vector of email.

Organizations need to evolve quickly or risk becoming obsolete.

There are two pieces to this puzzle. The first is the move to HRM that proactively manages human risk through hyper-personalized controls and guidance when it's needed most, transforming people into active participants in their organization's security.

The second shift is to train AI agents to act securely as they become more active members of the workforce.

It's a question of "when" not "if" this approach should be taken — and the early adopters will elevate themselves above those who delay.

About KnowBe4

KnowBe4 empowers workforces to make smarter security decisions every day. Trusted by over 70,000 organizations worldwide, KnowBe4 helps to strengthen security culture and manage human risk.

KnowBe4 offers a comprehensive AI-driven 'best-of-suite' platform for Human Risk Management, creating an adaptive defense layer that fortifies user behavior against the latest cybersecurity threats. The HRM+ platform includes modules for awareness & compliance training, cloud email security, real-time coaching, crowdsourced anti-phishing, AI Defense Agents, and more. KnowBe4 provides security training for both humans and AI agents, recognizing that as AI becomes more integrated into business operations, workforce trust management becomes an essential protection layer.

By training both humans and AI agents to recognize and respond appropriately to security risks, KnowBe4 ensures comprehensive defense strategies across an organization's entire workforce.

For more information, please visit www.KnowBe4.com.

Methodology

The data in this report is compiled from an independent survey conducted by Arlington Research of 700 global cybersecurity leaders and 3,500 global employees with no responsibility for cybersecurity.

knowbe4

KnowBe4, Inc. | 33 N Garden Ave, Suite 1200, Clearwater, FL 33755
855-KNOWBE4 (566-9234) | www.KnowBe4.com | Sales@KnowBe4.com