Osterman Research
WHITE PAPER

White Paper by Osterman Research
Published **January 2026**
Commissioned by **IRONSCALES**

Rebuilding Trust in Digital Communications

Executive summary

Current and emerging types of cyberattacks are undermining trust in digital communication channels and verified identities. New types of attacks that undermine trust are facilitated by AI, driving high levels of concern across organizations. But even existing attack types that we have already wrestled with for decades and would prefer to consider solved—such as phishing and business email compromise—are being **supercharged** by AI. The common culprit in new generations of trust-destroying attacks is AI, driving malicious innovations in deepfakes, synthetic media, impersonation, and phishing to compromise credentials.

Organizations can't afford to trust that existing security solutions are up to the task of detecting, stopping, and mitigating these new attacks. Training approaches must be revisited to cultivate awareness of new deepfake attacks. Technical cybersecurity protections, likewise, must be bolstered by investing in solutions that address new and emerging attack quadrants.

KEY TAKEAWAYS

The key takeaways from this research are:

- **Higher threat actor interest in attacking; elevated threat levels; almost everyone has been compromised**
  82% of organizations have seen higher interest from threat actors in compromising digital communications over the past year. Threat levels of AI-infused phishing attacks, vendor impersonation, and deepfake audio have increased most significantly. 88% of organizations have suffered at least one security incident that undermined trust in the past 12 months.
- **Employees aren't ready and many attacks are still immature**
  51% of organizations are highly concerned that employees aren't ready to defend against attacks that weaponize trust, and overwhelmingly so for employee groups viewed as high-priority attack targets by threat actors. Deepfake audio and video attacks are accelerating rapidly, with threat levels increasing significantly over the past year. While many organizations have begun preparing defenses, 60% lack confidence in their ability to counter these attacks effectively—a dangerous gap as threat actors continue to refine their capabilities.
- **Data breach risk is higher, efficacy of core activities lower**
  55% of respondents say that failing to counter attacks that weaponize trust significantly increases the likelihood of a data breach. Several additional costs follow closely behind—hits to employee productivity and workflow efficiency, a reduced ability to engage with customers, and declining market capitalization.
- **Strengthen, re-platform, and/or build your own next generation of protections**
  Organizations are re-evaluating how they assess security tools to safeguard digital communication channels and identities, given high levels of concern about threat actors' access to advanced capabilities for new attack types. Best-in-class point solutions, complete replacement, and build-your-own avenues are all under consideration.

> ***AI is supercharging existing attack types we would prefer to consider solved—and destroying trust in the process.***

ABOUT THIS WHITE PAPER

The survey and white paper were commissioned by IRONSCALES. Information about IRONSCALES and details on the survey methodology are provided at the end of the paper.

Undermining trust in digital communication channels and verified identities

Current and emerging types of cyberattacks are undermining trust in digital communication channels and verified identities. The presence of fake messages and fake participants in digital communication channels such as email, collaboration spaces (e.g., Microsoft Teams, Slack), and online meeting services (e.g., Teams, Zoom, Google Meet, Webex) results in employees and executives constantly questioning whether what is being said or presented is real.

The threat to trust is the same for identities. For any verified identity—meaning the password was correct and any multi-factor authentication (MFA) requirements were met—how does one know if it's an employee on the other end or a threat actor who has compromised the account and its MFA protections?

The net result is an undermining of trust.

Consider these four types of attacks that are illustrative of undermining trust:

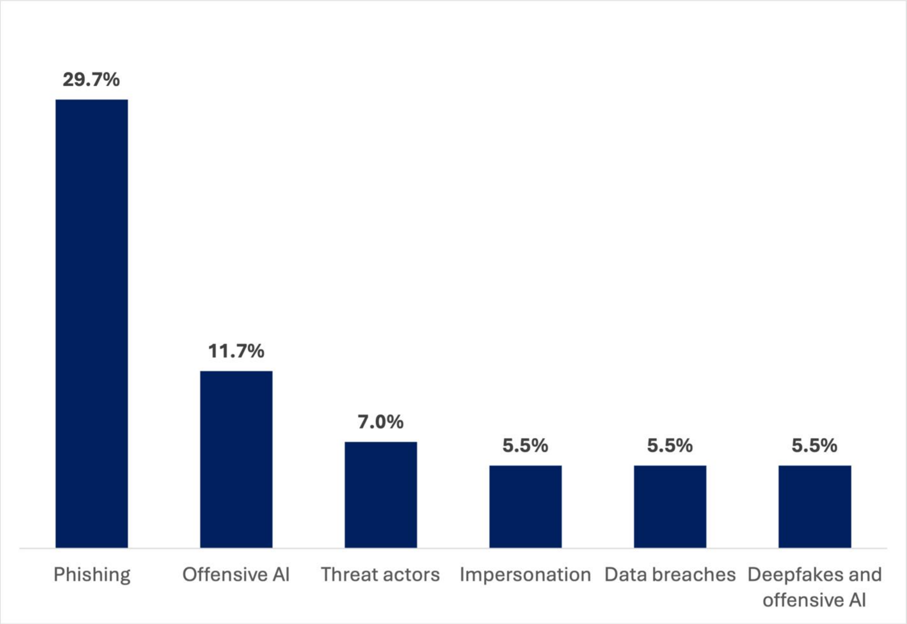
- **Buying compromised credentials**
  Threat actors are buying compromised credentials for employees from dark web forums. For threat actors, this gives a valid access path into an organization's systems without having to hack their way in. For organizations, it amplifies the importance of effective pre-access authorization checks, so that invalid use of valid credentials by threat actors can be detected, blocked, and mitigated. Many organizations are failing to do so.
- **Vendor masquerading**
  Threat actors are masquerading as trusted vendors to steal funds or information from organizations, for example, through a business email compromise (BEC) scam. BEC has proven to be a consistently profitable ruse, since such scams have repeatedly been one of the most expensive loss categories in the annual FBI report on internet crimes.[1]
- **Deepfake audio**
  Threat actors are using deepfake audio to trick employees into taking actions that aren't in their best interest, such as authorizing a payment because they believed the voice on the other end of the phone or the voice mail was their boss or a high-ranking executive. Deepfake audio can be created for any voice based on just a few seconds[2] of a valid recording—something that's easy to get for executives due to their participation in conference calls, investor summits, and social media posts.
- **Hiring fraud**
  Threat actors are attempting to get hired into positions with high-privilege access to systems and data, thus compromising organizations initially through hiring fraud. This strategy has been embraced by North Korean operatives with particular success,[3] often supported by malicious use of AI models from leading AI vendors, e.g., OpenAI[4] and Anthropic.

The presence of fake messages and fake participants in digital communication channels results in employees and executives constantly questioning whether what is being said or presented is real.

We asked survey respondents what they saw as the most significant current threat to the trustworthiness of digital communication channels and verified identities at their organization. After coding and grouping the open-ended responses, 65% of respondents gave one of six answers:

- Phishing, including specific mention of identity spoofing, impersonation, or social engineering. 30% of respondents said this was the most significant threat—almost three times as high as any of the five answers below.
- Offensive AI (not explicitly including deepfakes), with an emphasis on advanced AI tools for driving cyberattacks.
- The activities of threat actors—a generalized, high-level concern among respondents about everything that threat actors are doing to undermine the trustworthiness of digital communication channels and verified identities.
- Data breaches of various types of sensitive data, such as customer data and identity verification data. These can be weaponized for second-order attacks.
- Impersonation of employees or executives used for gaining unauthorized access to sensitive data.
- Deepfakes, AI, and synthetic media for tricking people using faked voice and video streams.

Figure 1
Threats to the trustworthiness of digital communication channels and verified identities
Percentage of respondents



Source: Osterman Research (2026)

Phishing (including identity spoofing, impersonation, and social engineering) is the most significant threat to trust in digital communication channels.

Threat actors undermining trust—the past 12 months

Over the past 12 months, the organizations in this research have seen an increase in activities to undermine trust, heightened threat levels across multiple types of attacks, and a high rate of security incidents.

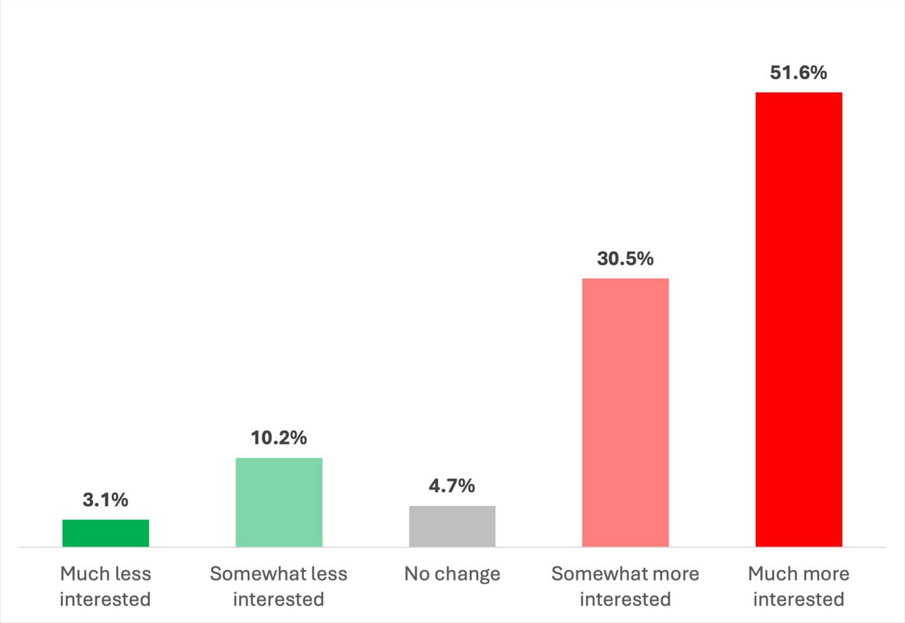
THREAT ACTORS ARE MORE ACTIVE

Threat actors have stepped up their attempts to exploit trusted digital communication channels and verified identities. 82% of the organizations in this research say that cybercriminals have shown a heightened interest in exploiting these trusted entities over the past 12 months. More than half—51.6%—peg this interest at the highest level.

See Figure 2.

51.6% of organizations say that cybercriminals have shown the highest level of interest in exploiting trusted digital communication channels and verified identities.

Figure 2
Change in threat actor interest in exploiting trust in digital communications over the past 12 months
Percentage of respondents

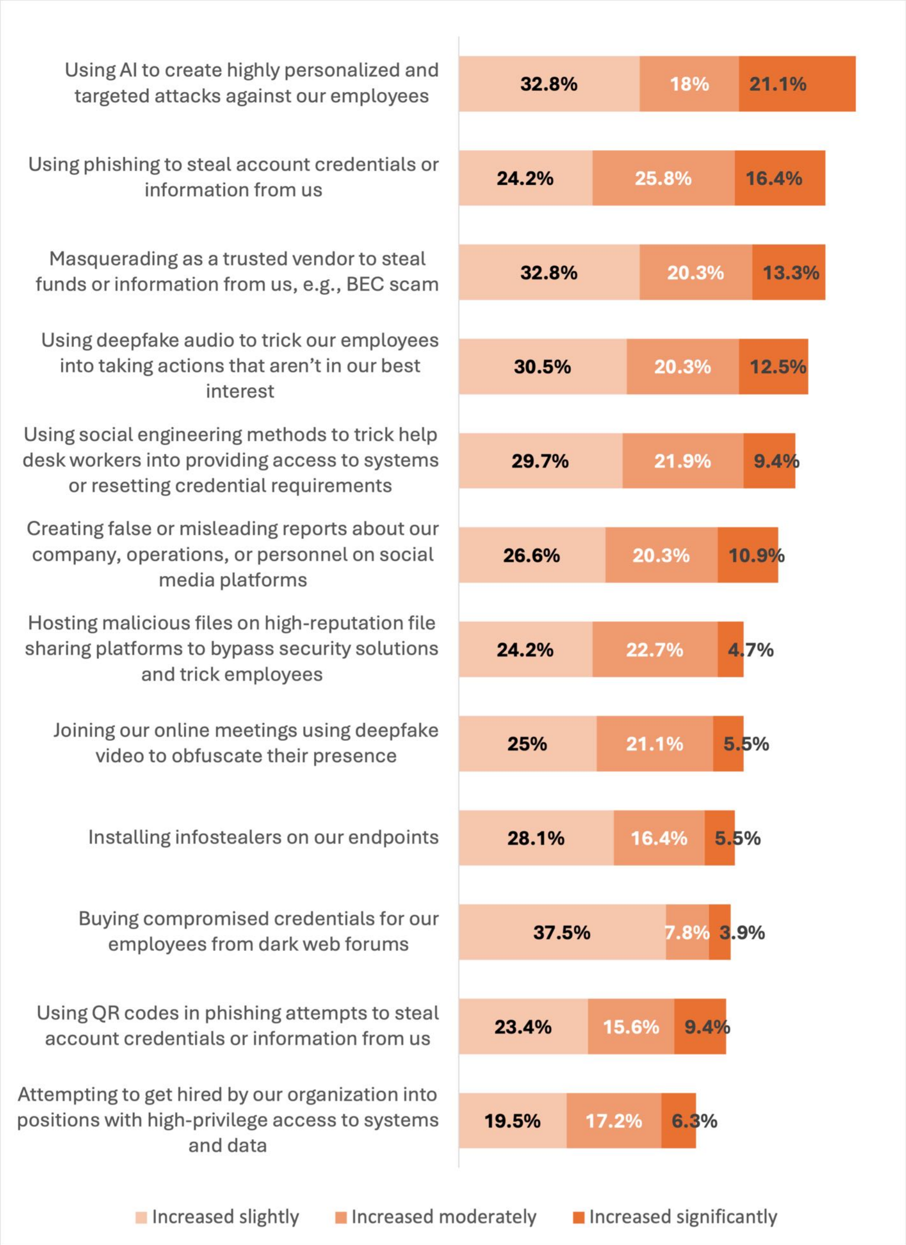


Source: Osterman Research (2026)

THREAT LEVELS ARE UP

Over half of respondents say that threat levels of the 12 attack types we explored in this research have increased over the past year, with the most significant increases seen for threat actors using AI to create highly personalized and targeted attacks against employees (21.1%), phishing to steal account credentials or information (16.4%), and attacks where a threat actor masqueraded as a trusted vendor to steal funds or information (13.3%). Phishing and BEC attacks are long-running attack types, and partly due to the use of AI to enhance these attacks, they are continuing to get worse. The threat level of deepfake audio is in fourth place. See Figure 3.

Figure 3
Change in threat level of actions taken by threat actors over the past 12 months
Percentage of respondents



Source: Osterman Research (2026)

Phishing and BEC attacks continue to get worse, partly due to threat actors using AI to enhance these attacks.

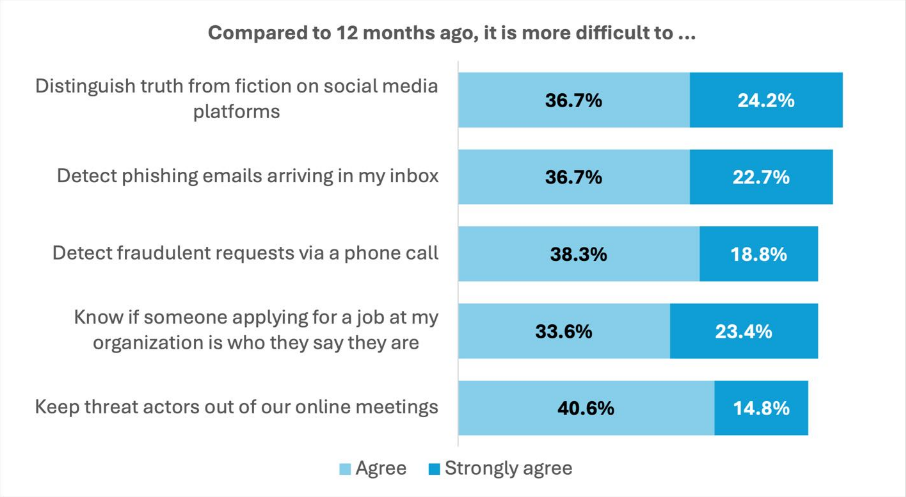
COMMON DIGITAL COMMUNICATION CHANNELS HAVE BEEN COMPROMISED

Respondents acknowledge that it is more difficult to trust what is shown or presented in common digital communication channels compared to 12 months ago. On average, 58% of respondents indicate that five detection tests or checks have become more difficult. The highest level of difficulty is experienced for distinguishing truth from fiction on social media platforms (24.2%), detecting hiring fraud (23.4%), and detecting phishing emails (22.7%). See Figure 4. Consider:

- **True or false?** Apps from OpenAI, Google, and Meta enable realistic short form videos to be created from a written prompt or photograph.[5] AI companies are actively releasing technologies that enable the creation of imaginary situations which are presented as real—something threat actors will rapidly embrace because it supports and enables their threat playbook.
- **Hiring fraud**. Several trends have made hiring fraud, particularly of IT workers, an easier and profitable ruse for threat actors. Remote work is normalized. Drawing on global hiring pools has become common practice. Threat actors utilize VPNs to obfuscate their location in combination with on-the-ground domestic operatives running laptop farms in their homes.[6] Finally, corporate use of AI models is increasingly expected of employees for performing work, developing code, and answering technical questions, and this aggressive corporate adoption makes it easier for imposters to use AI-generated work product without being detected as a fraudulent hire.
- **New generations of phishing**. Phishing emails are significantly more advanced due to threat actors using AI. Early forays into the use of malicious AI models (e.g., WormGPT) appear to have been replaced by threat actors learning how to obfuscate malicious intent and bypass safety guardrails when using AI models from Google, OpenAI, and Anthropic, among others. The use of AI for generating targeted phishing has eliminated the traditional indicators that people and systems were taught to look for when evaluating messages. The use of compromised identities for internal phishing or third-party attacks, likewise, has made it more difficult to know whether a message is from the person normally using an account—or a threat actor who has compromised it.

Distinguishing truth from fiction, detecting hiring fraud, and detecting phishing emails is more difficult than a year ago.

Figure 4
Difficulty of detecting threat actor activity
Percentage of respondents



Source: Osterman Research (2026)

SECURITY INCIDENTS HAVE UNDERMINED TRUST MULTIPLE TIMES IN THE PAST 12 MONTHS

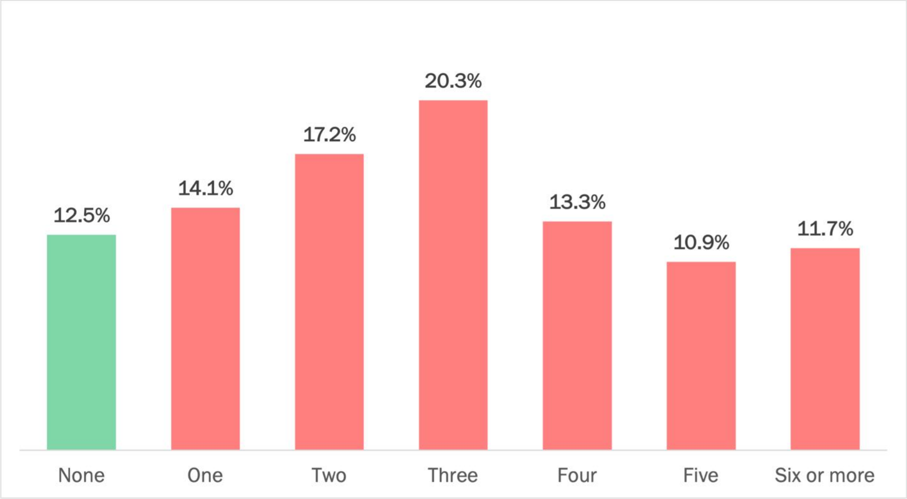
During the past 12 months, 87.5% of the organizations in this research have experienced at least one security incident that undermined trust in digital communication channels and verified identities.

The four most common incident types were:

- Threat actors bypassed security solutions and tricked employees into downloading a malicious file from a high-reputation file sharing platform (35.2% of organizations).
- Threat actors masqueraded as a trusted vendor and stole funds or information, e.g., through a business email compromise scam (34.4%).
- Threat actors used compromised credentials to gain access to systems or data (32.8%). Authorization systems allowed the authentication request, not realizing a threat actor was impersonating an employee.
- Threat actors used artificial intelligence to create highly personalized and targeted attacks against employees (31.3%).

See Figure 5.

Figure 5
Number of security incidents that undermined trust over the past 12 months
Percentage of respondents



Source: Osterman Research (2026)

Security incident types reflected in Figure 5 happened at least once at the organizations in this research. They are likely to have occurred multiple times, however, because if security controls are ineffective at stopping a particular threat type, it is likely such threats will slip through undetected or unreported until better security controls are implemented.

87.5% of the organizations in this research have experienced at least one security incident that undermined trust in digital communication channels and verified identities during the past 12 months.

Organizations are ill-prepared for attacks that undermine trust

As threat actors ramp up attacks that weaponize trust, many organizations find themselves ill-prepared to detect and defend against such attacks.

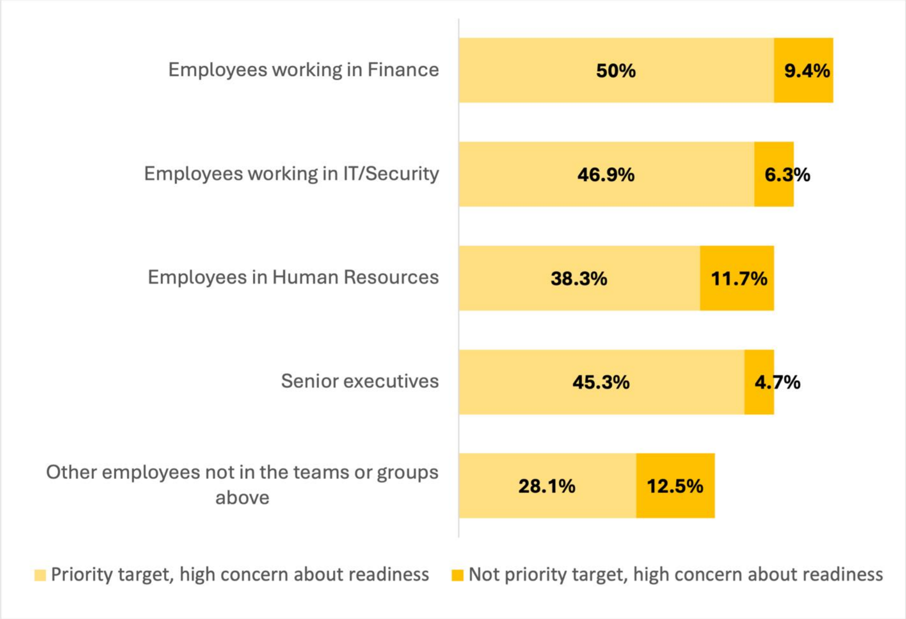
KEY EMPLOYEE GROUPS ARE NOT READY TO DEFEND AGAINST ATTACKS

On average, 50.6% of organizations indicate high levels of concern about the readiness of employees to defend against attacks that weaponize trust, for both employees in general and employees in groups or roles of specific interest to threat actors. Levels of concern are highest for employees working in finance roles—a group that is of particular interest to threat actors since finance employees have access to payment systems. The second highest level of concern is for employees in IT and security roles, because although these employees don't have access to payment systems directly, they control access to systems, data, and privileges that can be compromised either to gain access to funds directly (through privilege escalation) or through extortion schemes after breaching data.

Figure 6 correlates where high levels of concern about readiness ("highly concerned" or "extremely concerned") align with whether the respondent believes the employee groupings are a high target priority for threat actors ("high priority" or "extreme priority") or not. This correlation is for the respondent's organization, not for all organizations in general. On average, over 80% of the elevated concern level exists where the respondent also sees the employee groups under high priority attack from threat actors—an alarming combination.

The readiness of employees working in finance, IT, and security roles to defend against attacks that weaponize trust is of high concern.

Figure 6
Attack priority and readiness concern for employee groupings
Percentage of respondents



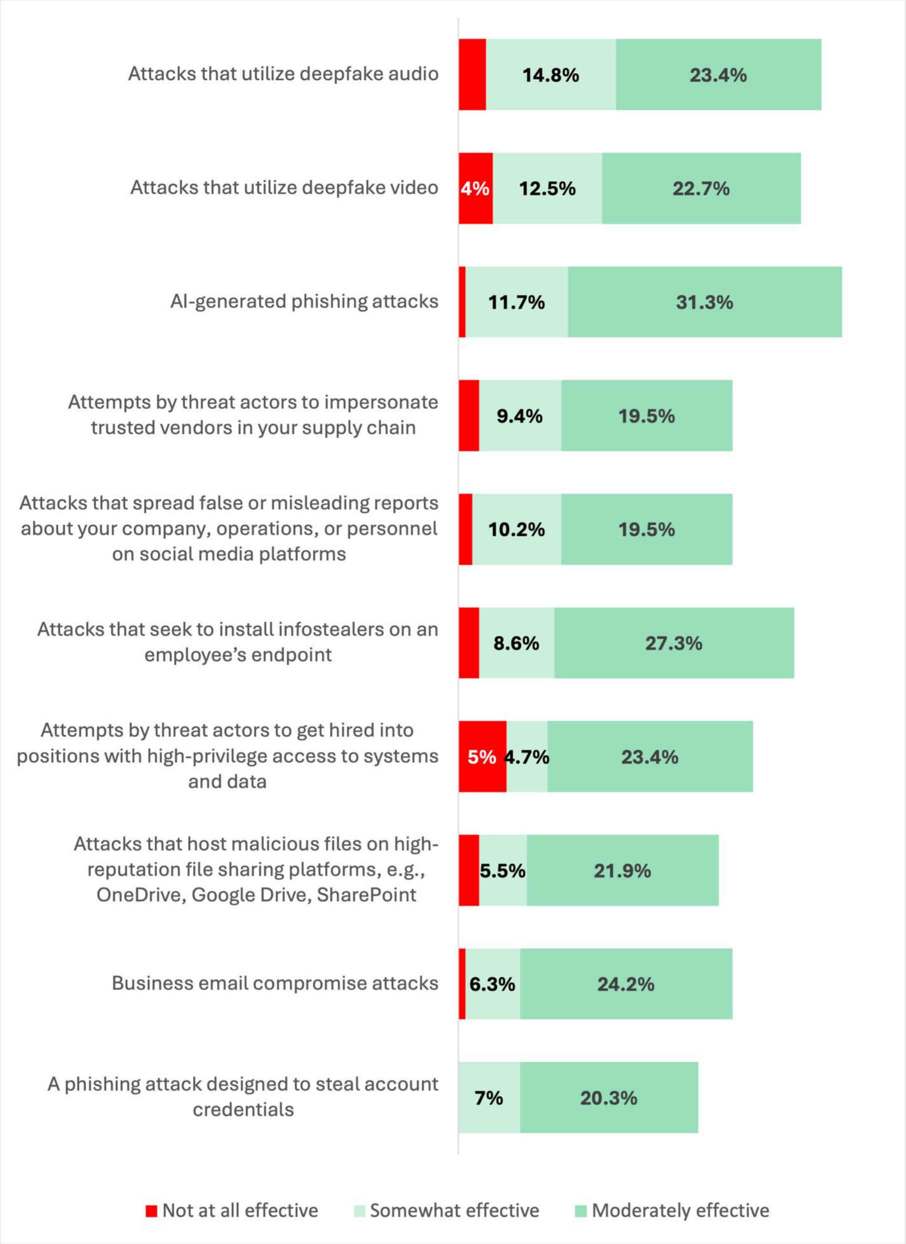
Source: Osterman Research (2026)

TRAINING APPROACHES AREN'T ADDRESSING EVOLVING THREATS

Current training approaches for preparing employees to detect attacks that weaponize trust are proving ineffective for many organizations. Training on detecting attacks utilizing deepfake audio and video are particularly ineffective, and detecting AI-generated phishing is in third place. The attack type with the highest result for "not at all effective" is hiring fraud for positions with high-privilege access to systems and data (5%).

See Figure 7.

Figure 7
Low effectiveness of current employee training approaches for detecting attacks
Percentage of respondents sorted by the sum of "not at all effective" and "somewhat effective"



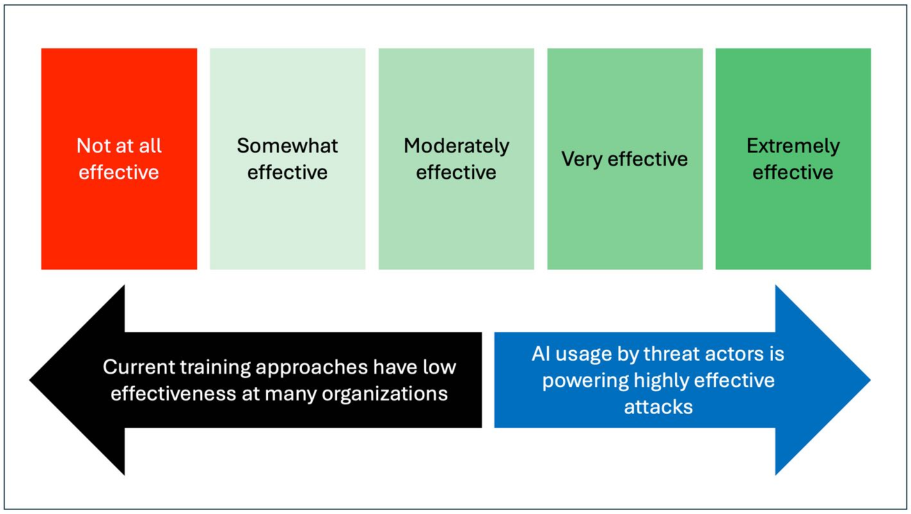
Source: Osterman Research (2026)

> ***Many organizations are ineffective at preparing employees to detect attacks that weaponize trust.***

For many organizations, current training approaches for detecting attacks that weaponize trust trend toward the lower end of the effectiveness continuum. By contrast, threat actors are using AI, deepfakes, synthetic media, and other advanced malicious innovations to create new generations of attacks which trend towards the higher end of the effectiveness continuum. The mismatch between the two creates significant openings for threat actors to target organizations with attacks that neither they nor their employees are prepared to detect, counter, and rebuff.

See Figure 8.

Figure 8
Threat actors versus organizations: mismatched realities


Not at all effective
Somewhat effective
Moderately effective
Very effective
Extremely effective
Current training approaches have low effectiveness at many organizations
AI usage by threat actors is powering highly effective attacks

Source: Osterman Research (2026)

THREAT ACTORS ARE JUST GETTING STARTED

Respondents believe that threat actors are in the early stages of several attack types, with attacks using deepfake audio, deepfake video, and AI-generated phishing viewed as being immature and emergent by half of respondents. See Figure 9. In other words, we haven’t seen anything yet.

An average of 13% of respondents say the other attack types in Figure 9 are immature and emergent, including attacks like BEC and vendor impersonation. While on first glance that may seem like it couldn’t be true, the reality is that it’s deeply concerning. Threat actors are embracing new threat innovations and combining deepfake technology, synthetic media, offensive AI, and cybercrime-as-a-service capabilities to design new generations of attack types.

Known attack types including phishing and business email compromise have existed for several decades already, but AI has fundamentally reset their maturity curve. **The BEC attacks of 2025 bear little resemblance to the BEC attacks of 2020**. Today’s BEC attacks are hyper-personalized, multi-channel, and can be launched autonomously at scale. In effect, we’re facing an immature and emergent threat that just happens to use a familiar attack category. When 2030 arrives, it’s likely we’ll look back and comment on how significantly attacks have changed since this research was conducted.

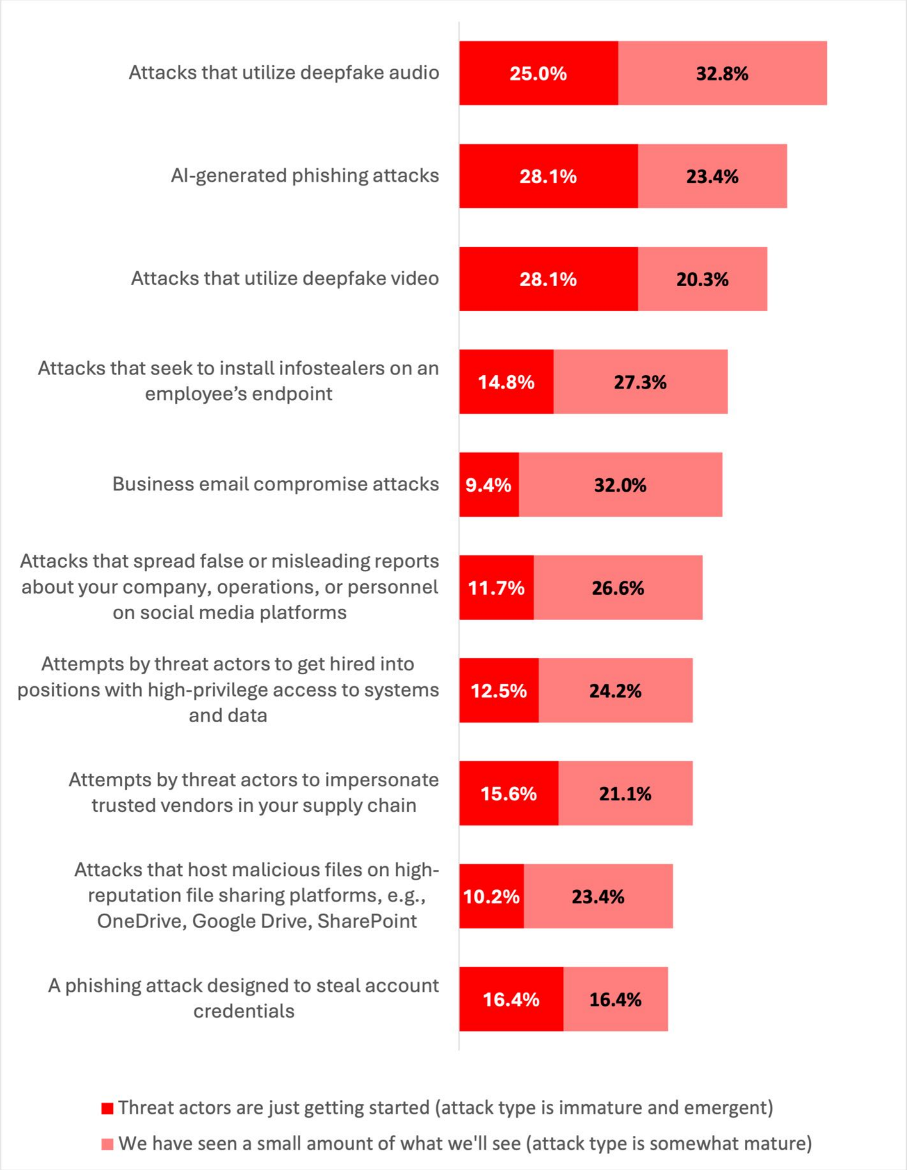
> ***Threat actors are embracing new threat innovations and combining deepfake technology, synthetic media, offensive AI, and cybercrime-as-a-service capabilities to design new generations of attack types.***

In summary, the threat curve just got reset. Even "solved" attack types like phishing and BEC have become immature again because:

- The barrier to entry collapsed, since anyone can now craft attacks in perfect English, German, Spanish, or any other language.
- The sophistication ceiling has been raised due to hyper-personalization at scale informed by the aggregation of target profiles from data breach records.
- The attack surface exploded—it's now multi-channel, multi-modal, and autonomous.

See Figure 9.

Figure 9
Assessing the maturity and emergence of sophisticated attack types
Percentage of respondents



Source: Osterman Research (2026)

AI has reset the threat curve for "solved" attack types like phishing and BEC.

HIGH CONCERN ABOUT DEEPFAKE IMPERSONATION, PHISHING, AND OFFENSIVE AI OVER THE NEXT 12 MONTHS

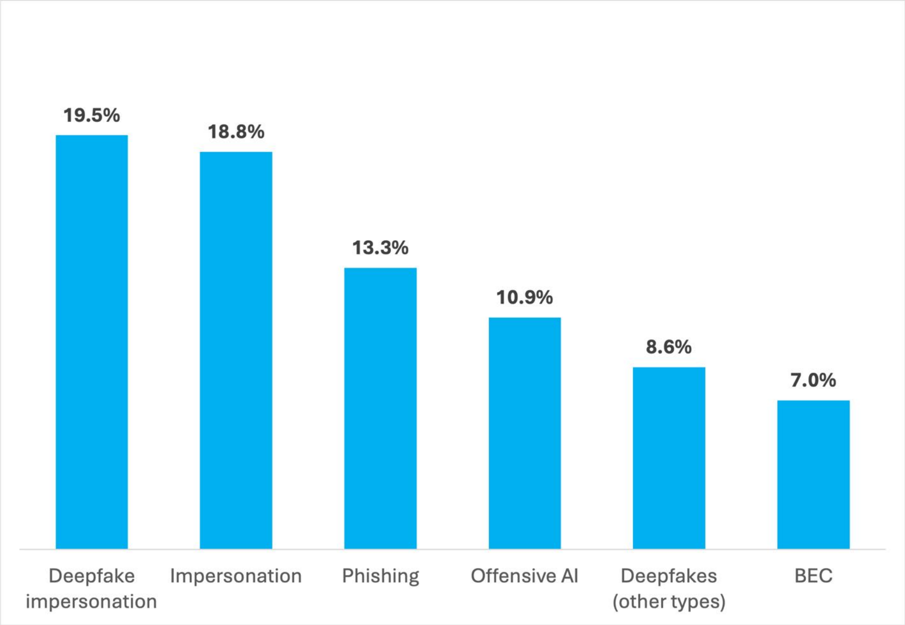
We asked survey respondents what specific type of impersonation or social engineering they were most concerned about over the next 12 months. After coding and grouping the open-ended responses, 78% of respondents gave one of five answers:

- Deepfake impersonation, most commonly of executives and for uses in AI-generated phishing and vishing attacks.
- Other types of impersonation not specifically noted as deepfakes, most commonly for breaching sensitive information and the impersonation of employees or executives.
- Phishing. Several respondents noted phishing in conjunction with offensive AI and the increasing sophistication of phishing attacks.
- Offensive AI more generally, for increasing the scale and speed of attacks, bypassing traditional protections, and resulting in data breaches the organization can't stop.
- Business email compromise attacks resulting from social engineering and impersonation scams.

See Figure 10.

Deepfake impersonation, most commonly of executives and for use in AI-generated phishing and vishing attacks, is of high concern in the next 12 months.

Figure 10
Concern for specific types of impersonation or social engineering over the next 12 months
Percentage of respondents



Source: Osterman Research (2026)

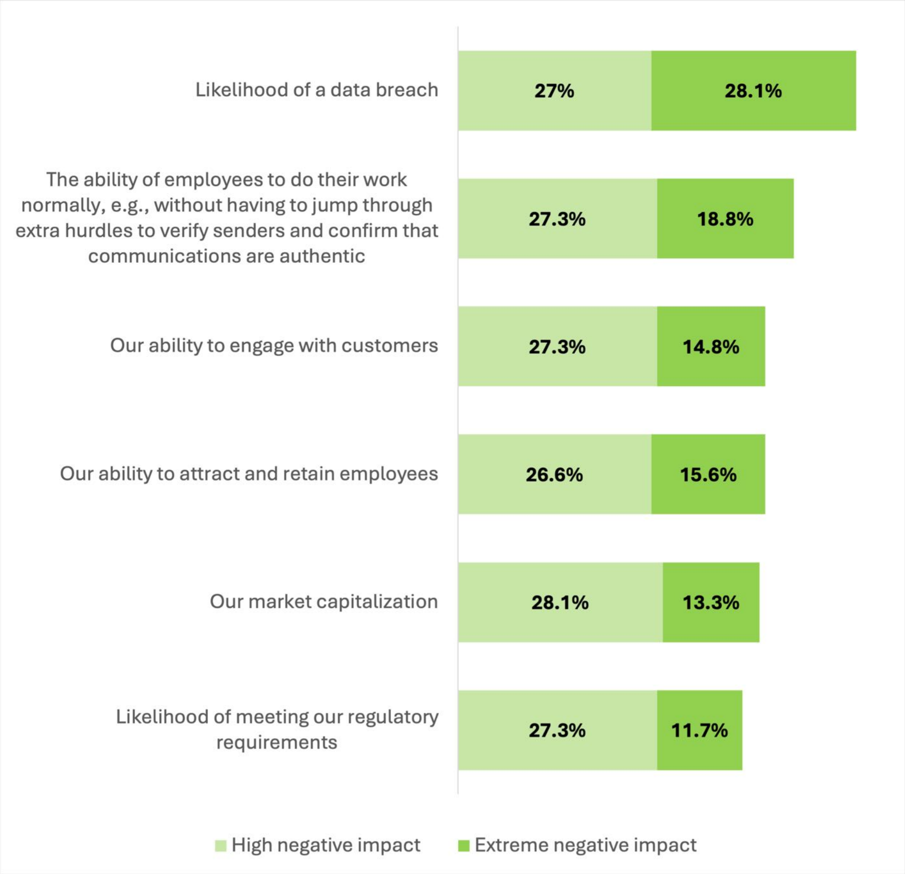
The costs of compromised trust

Organizations will incur negative costs if they don't successfully defend against attacks that weaponize trust over the next 12 months. For the organizations in this research, the increased likelihood of a data breach is viewed as having the greatest negative impact. By a significant margin, this negative impact has the highest "extreme" rating of the six impacts we asked about. Threat actors have data in their sights for use in extortion campaigns. Increasingly sophisticated deepfake and AI-powered phishing attacks represent the next wave of how threat actors are plotting to make this happen.

Several other negative impacts follow closely behind in the ratings if attacks are not addressed—such as hits to employee productivity and workflow efficiency; a reduced ability to engage with customers and attract and retain employees; and dampened market capitalization.

See Figure 11.

Figure 11
Impacts of failing to defend against emerging attacks that weaponize trust
Percentage of respondents



Source: Osterman Research (2026)

> ***The increased likelihood of a data breach is viewed as having the greatest negative impact for organizations that don't successfully defend against attacks that weaponize trust.***

Rebuilding trust—a roadmap

The research data is clear: new generations of attacks that undermine trust in digital communication channels and verified identities have got many organizations highly concerned. Based on the data from this research, this section outlines a roadmap for rebuilding trust.

DEFEND AGAINST DEEPFAKE AUDIO AND VIDEO ATTACKS

AI has unleashed a new wave of sophisticated attack types, with deepfake audio and video among the most concerning. Relying on individuals to spot abnormalities in video streams, recognize imposters when they join online meetings, and discern the difference between their boss's voice and a synthetic version is not a route to success. While cultivating an awareness of the possibility of such scams through targeted cybersecurity awareness training and simulations is important, relying on individuals to detect and stop such attacks without any technological support is a recipe for failure.

Existing vendors and startups are tackling deepfake impersonation scams with technology that analyzes voice and video streams for anomalous digital patterns and internal inconsistencies. These signals are analyzed in conjunction with normal device profile characteristics (e.g., Jim always uses an iPhone), geographical indicators (e.g., Jim doesn't travel out of his home country), and network connectivity attributes for each given user to quickly identify where deviations from baseline activity indicate likely fraudulent activity.

When identified, individual detection signals can be aggregated to drive an automated action of blocking, removing, or stopping a deepfake attack from continuing. In the case of online meetings, for example, a deepfake participant can be automatically removed from the meeting, or in certain cases defined in advance by policy and allowed to remain in the meeting for observational and forensics purposes.

Next action: Invest in solutions that detect deepfake audio and video attacks.

> ***Relying on individuals to detect and stop deepfake audio and video attacks without any technological support is a recipe for failure.***

SAY GOODBYE TO LEGACY EMAIL PROTECTIONS AND HELLO TO THE NEXT GENERATION

While deepfake audio and video represent the flamboyant application of AI to cyberthreat design, its use to enhance phishing campaigns is the quiet and deadly one. New generations of phishing attacks—including BEC campaigns and vendor impersonation scams—eliminate the historical telltale signs of a phishing attack. Spelling and grammar are perfect. Valid accounts have been compromised rather than using Gmail accounts with the right name. Tone- and style-perfect messages blend seamlessly into current email threads.

Legacy email protections won't help organizations defend against AI-powered phishing attacks. Secure email gateways (SEGs) and email security solutions designed to look for malicious links, weaponized attachments, and account impersonations are too blunt an instrument to recognize the subtle indicators of modern and still emerging AI-powered attacks. Significantly greater precision is required to detect, stop, and remediate the next generation of phishing attacks, and only AI-enabled email security defenses are up to the task.

Next action: Stop relying on legacy SEGs and email security solutions developed for a threat environment that no longer exists.

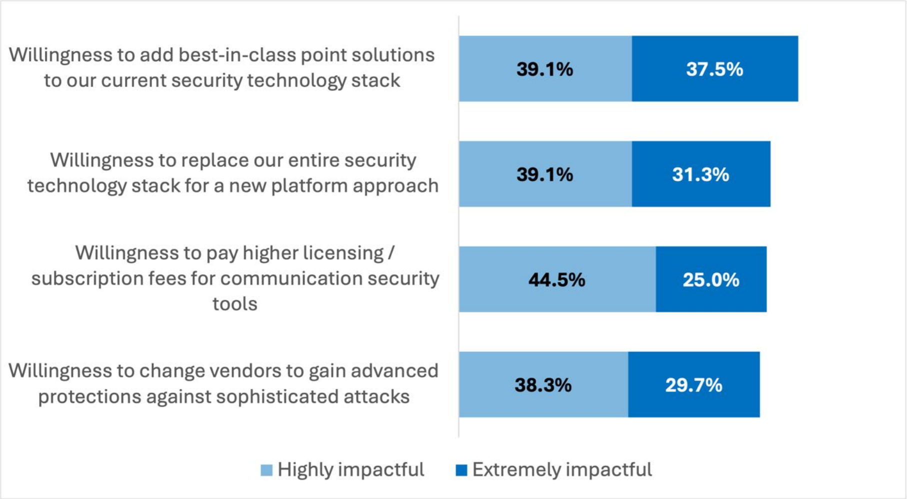
IF NECESSARY, TAKE DRASTIC ACTION TO SAFEGUARD TRUST

65% of respondents are "very concerned" or "extremely concerned" that threat actors increasingly have access to highly advanced capabilities for use in emerging sophisticated attack types, such as deepfake video calls. Given all the evidence we've looked at so far in this research, the others likely should be too.

As a consequence, organizations are evaluating communication security tools differently. The willingness to add best-in-class point solutions is the decision that's impacted the most, followed by an express willingness to shift to an entirely new platform approach to security—replacing everything currently in place with a new approach that has demonstrably better capabilities. Emerging sophisticated attack types that undermine trust and threaten the integrity of digital communication channels and identities are of such high concern that organizations signal a willingness to take drastic action to protect themselves, the data they're entrusted with, and the customers they serve.

See Figure 12.

Figure 12
Impact of concern over advanced attack capabilities on security decisions
Percentage of respondents



Source: Osterman Research (2026)

Take drastic action to protect digital communication channels and verified identities from attacks that weaponize trust if current defenses are ineffective.

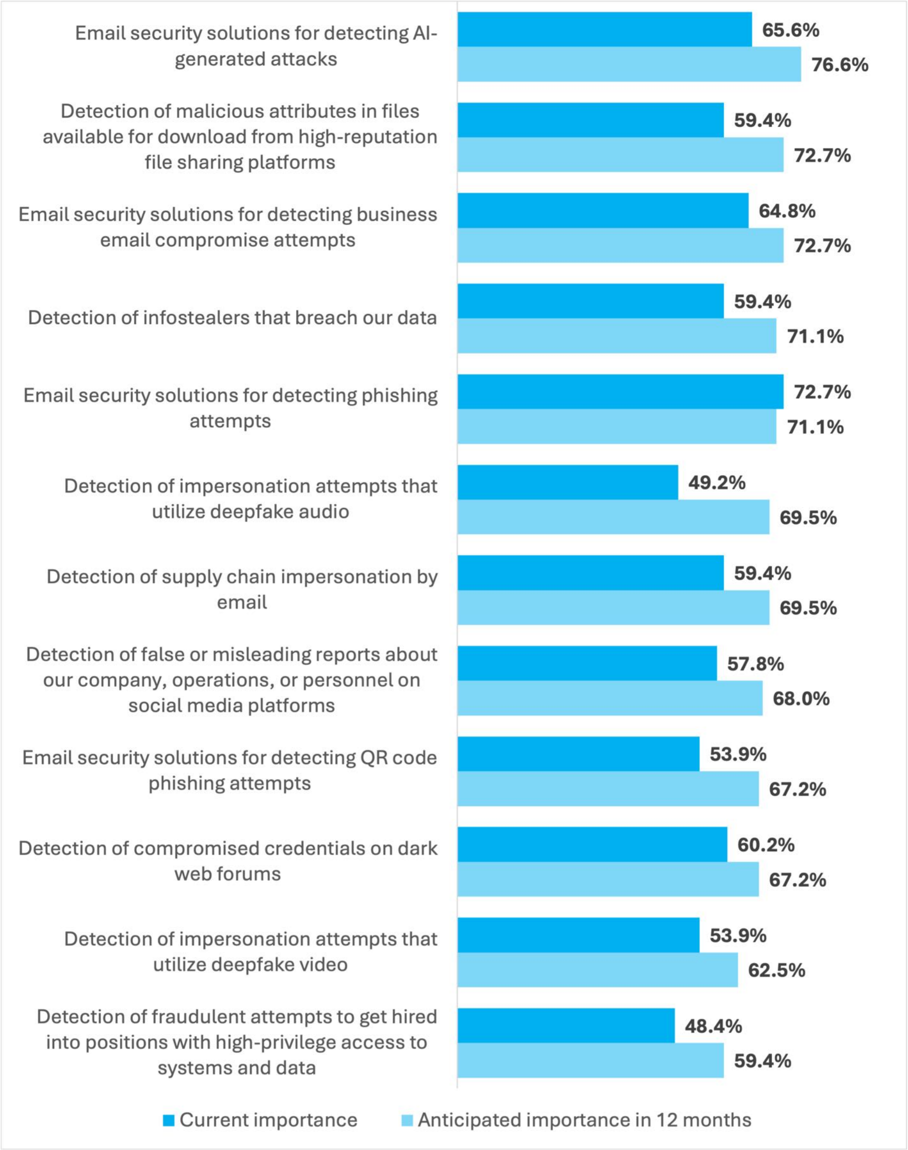
Next action: If current defenses and security solutions aren't up to it, take drastic action to protect digital communication channels and verified identities from attacks that weaponize trust.

STRENGTHEN DEFENSIVE POSTURE

The importance of defensive security technologies or strategies to the organizations in this research is anticipated to increase over the next 12 months. The largest increase in importance is for detecting deepfake audio impersonation attacks, followed by QR code phishing attempts and hiring fraud. The average increase in importance across the 12 technologies or strategies we asked about is 17.4%.

Only one of the technologies or strategies has an anticipated decline in importance—that being email security solutions for detecting phishing attempts (2.2% reduction). However, with the changing nature of phishing due to AI and deepfake technology, this reduction should be compared with the increase in importance of email security solutions for detecting AI-generated attacks—which has the highest level of importance in 12 months' time. See Figure 13.

Figure 13
Importance of defensive security technologies and strategies
Percentage of respondents indicating "very important" or "extremely important"



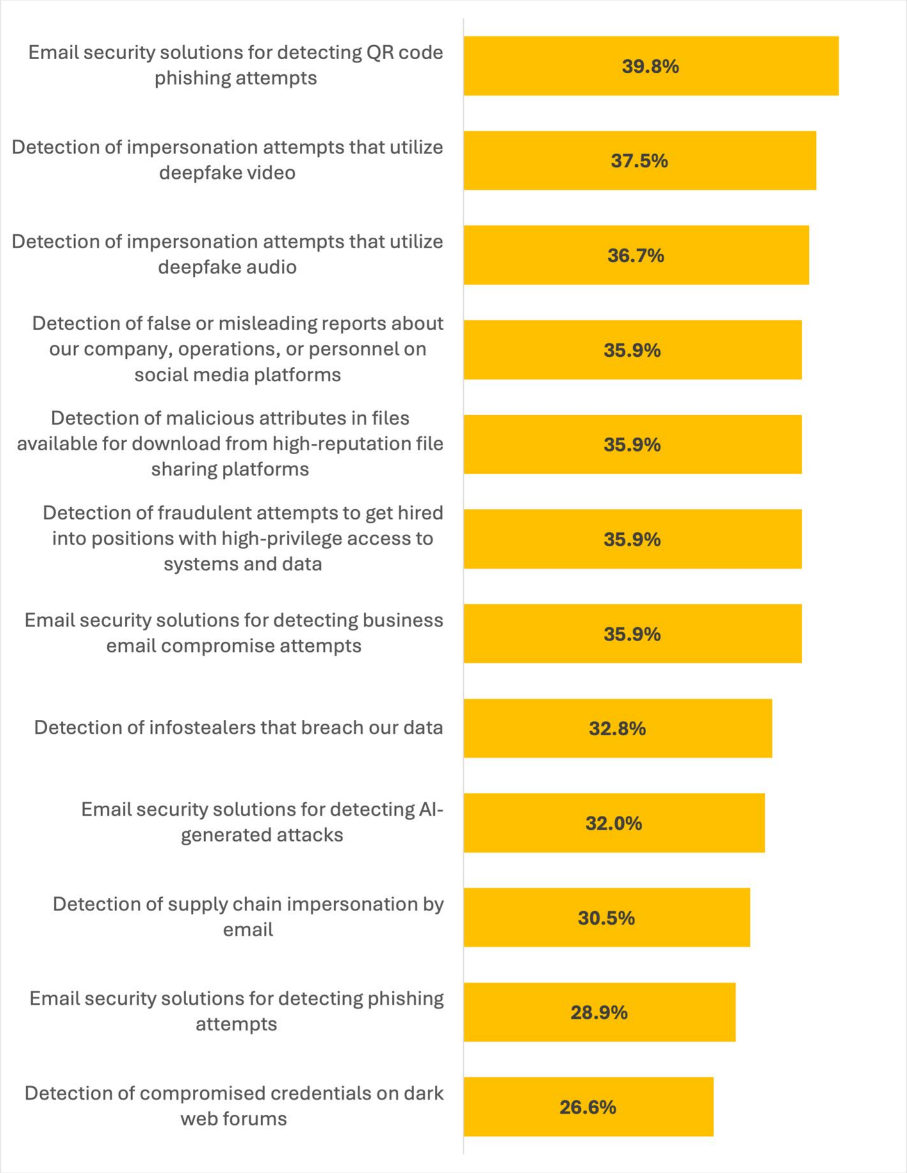
Source: Osterman Research (2026)

The importance of detecting impersonation attacks that utilize deepfake audio is anticipated to increase the most over the next 12 months.

On average, 34% of the organizations in this research indicate that the importance of the defensive security technologies and strategies in Figure 14 will increase over the next 12 months. This reflects the proportion of organizations who see a stepwise increase of at least one importance rating over the next 12 months, for example from "very important" to "extremely important."

See Figure 14.

Figure 14
Importance of defensive security technologies and strategies
Percentage of respondents indicating importance is higher in 12 months compared to current level of importance



Source: Osterman Research (2026)

Next action: Assess the effectiveness of your current security technologies and strategies against the threat patterns seen at your organization and act accordingly. Strengthen what needs to be bolstered.

An average of 34% of organizations anticipate that the importance of their defensive posture will increase over the next 12 months.

BUILD YOUR OWN TECHNOLOGY STACK?

Less than a third of organizations in this research express an intent to build security technologies inhouse, rather than relying on commercially available offerings from security vendors. Building inhouse can offer a valid path for organizations with the appropriate skills and processes internally. It does provide a way for building a solution that is rightsized for the organization, with the specific capabilities required instead of purchasing, for example, a larger and more complicated commercial offering that doesn't offer value for money. On the other hand, it risks missing wider industry innovations due to the dedicated focus that security vendors bring to their solution development.

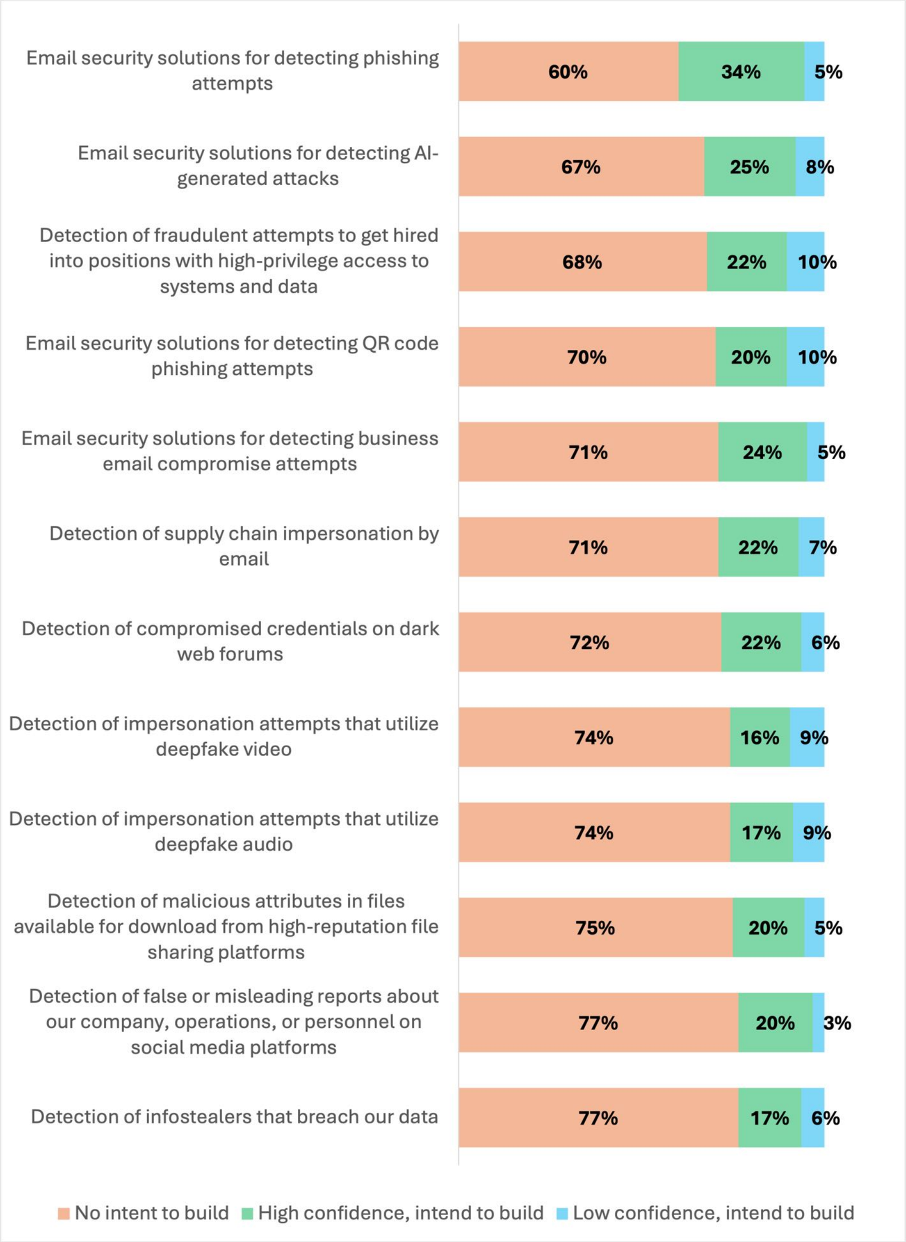
On average, 22% of the organizations in this research intend to build something inhouse while also having high confidence in commercially available offerings. The combination infers the intended development is more likely to be a product refinement, minor add-on, or specific capability that the commercial offering of choice doesn't address yet—or doesn't address to the level of maturity the organization seeks.

A much smaller set of organizations—7% on average—intend to build security technologies inhouse over the next 12 months while having low confidence in commercial offerings. This combination, by contrast, is more likely to reflect a substantial build process to address gaps that the organization believes they cannot address via commercial routes. Without year-on-year data due to this being our inaugural research program on this topic, how long this intent lasts remains to be seen.

See Figure 15.

> ***Less than a third of organizations intend to build security technologies inhouse, and most of these are likely to be pursuing a product refinement or minor add-on.***

Figure 15
Intent to build security technologies inhouse
Percentage of respondents



Source: Osterman Research (2026)

Building inhouse can offer a valid path for organizations with the appropriate skills and internal processes.

Next action: If your organization has the appropriate skills and processes and appropriate commercial offerings are not available or are unsuitable, consider building security technologies inhouse.

Conclusion

Trust is under attack as threat actors weaponize AI to create new ways of compromising digital communication channels, impersonating executives and vendors, and supercharging cyberattacks. Deepfake audio and video, along with AI-powered phishing, are prime methods that cause people to question the truth of what they see, hear, and read in online meetings, on the phone, and in their inbox. While it is still early days for some of these emergent attack types, organizations anticipate that both emergent and existing attacks still have substantial room to get significantly worse. It is critical that organizations urgently strengthen both technical and human protections against attacks that undermine trust.

> ***It is critical that organizations urgently strengthen both technical and human protections against attacks that undermine trust.***

About IRONSCALES

IRONSCALES is the leader in AI-powered email security, protecting over 17,000 global organizations from advanced phishing threats. As the pioneer of adaptive AI, we detect and remediate attacks like business email compromise (BEC), account takeovers (ATO), and deepfake attacks that other solutions miss. By combining the power of AI and continuous human insights, we safeguard inboxes, unburden IT teams, and turn employees into a vital part of cyber defense across enterprises and managed service providers.

IRONSCALES is headquartered in Atlanta, Georgia.

Visit www.ironscales.com or @IRONSCALES to learn more.

IRONSCALES
SAFER TOGETHER

www.ironscales.com

@IRONSCALES

Methodology

This white paper was commissioned by IRONSCALES and conducted by Osterman Research. One hundred twenty-eight (128) respondents who have direct responsibility for managing the cybersecurity posture at their organization were surveyed during September 10 to October 7, 2025. To qualify, respondents had to work at organizations with between 1,000 and 5,000 employees. All surveys were conducted in the United States. The survey was cross-industry, and no industries were excluded or restricted.

JOB ROLE

IT manager, director or VP	39.8%
CIO, or some other role that has this responsibility	15.6%
CISO, or some other role that has this responsibility	14.8%
CTO, or some other role that has this responsibility	14.8%
Security director or VP	14.8%

ORGANIZATION SIZE

1,000 to 1,999 employees	34.4%
2,000 to 4,000 employees	44.5%
More than 4,000 employees	21.1%

INDUSTRY

Information technology	16.4%
Industrials (manufacturing, construction, etc.)	15.6%
Retail or ecommerce	14.8%
Financial services	12.5%
Healthcare	8.6%
Professional services (law, consulting, etc.)	7.0%
Education	6.3%
Agriculture, forestry or mining	4.7%
Computer hardware or computer software	4.7%
Energy or utilities	3.1%
Transport or logistics	3.1%
Data infrastructure or telecom	1.6%
Life sciences or pharmaceuticals	0.8%
Media or creative industries	0.8%

Osterman Research does not provide legal advice. Nothing in this document constitutes legal advice, nor shall this document or any software product or other offering referenced herein serve as a substitute for the reader's compliance with any laws (including but not limited to any act, statute, regulation, rule, directive, administrative order, executive order, etc. (collectively, "Laws")) referenced in this document. If necessary, the reader should consult with competent legal counsel regarding any Laws referenced herein. Osterman Research makes no representation or warranty regarding the completeness or accuracy of the information contained in this document.

THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. ALL EXPRESS OR IMPLIED REPRESENTATIONS, CONDITIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE DETERMINED TO BE ILLEGAL.

[1] FBI, FBI Internet Crime Report 2024, April 2025, https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf

[2] John Hudson and Hannah Natanson, A Marco Rubio impostor is using AI voice to call high-level officials, July 2025, https://www.washingtonpost.com/national-security/2025/07/08/marco-rubio-ai-imposter-signal/

[3] CrowdStrike, CrowdStrike 2025 GlobalThreat Report, February 2025, https://www.crowdstrike.com/en-us/global-threat-report/

[4] OpenAI, Disrupting malicious uses of AI: June 2025, June 2025, https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-june-2025/

[5] Brian X. Chen, A.I. Video Generators Are Now So Good You Can No Longer Trust Your Eyes, October 2025, https://www.nytimes.com/2025/10/09/technology/personaltech/sora-ai-video-impact.html

[6] Robert McMillan, American Sentenced to 8½ Years in Prison for Helping North Koreans Get Jobs at Nike, Other U.S. Firms, July 2025, https://www.wsj.com/us-news/law/american-sentenced-to-8-years-in-prison-for-helping-north-koreans-get-jobs-at-nike-other-u-s-firms-d7de8be7