PHIN



2026 Whitepaper

# How AI is Changing Cyber Threats

## And What Your MSP Can Do About It



Learn more at
phinsecurity.com

# Table of Contents


03 *Overview*

04 *What Is Social Engineering, and How Has AI Changed It?*

06 *Deepfakes*

09 *Business Email Compromise*

12 *Five Red Flags of Modern Social Engineering*

14 *The Business Risks of AI-Powered Social Engineering*

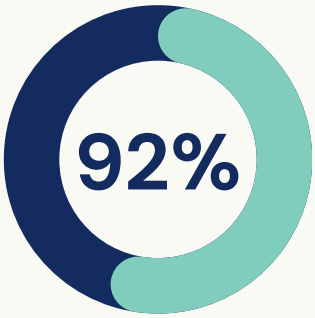
16 *How MSPs Can Build Stronger Defenses* Against AI

20 *Fighting AI-Powered Threats: People, Process, & Technology*

**Artificial intelligence (AI) has given cybercriminals new ways to manipulate people, impersonate trusted individuals, and create convincing scams at a scale that wasn't previously possible.** An attacker can generate a professional-looking email in seconds, clone an executive's voice, or create a realistic video conference featuring people who aren't actually there. And the numbers suggest this is becoming a serious problem.

Year-over-year increase in AI-enabled malicious breaches

**56%**




of organizations reporting AI-related breaches lacked proper AI access controls

Increase in reported consumer fraud losses from 2020 to 2025

**430%**

*Source: IBM, Cost of a Data Breach Report 2026*

*Source: FTC, 2026 congressional testimony*

These statistics highlight two important realities:

1. AI is creating new opportunities for attackers.
2. Many organizations are struggling to keep their security practices up to speed.

For managed service providers (MSPs), this creates a growing responsibility to help clients understand and defend against increasingly sophisticated threats.

**The key is combining proven security practices with modern technology that helps organizations detect, investigate, and respond to threats more effectively.**

While AI has changed how social engineering attacks are carried out, the fundamentals of preventing them haven't changed nearly as much.

# What Is Social Engineering, and How Has AI Changed It?

**Social engineering is the psychological manipulation of people** into revealing sensitive information, providing unauthorized access, transferring money, or taking other actions that benefit an attacker. Rather than exploiting a technical vulnerability, **social engineering exploits human behavior.**




of breaches involved social engineering in the past year, with an average breach cost of $5.23M.

*Source: IBM, Cost of a Data Breach Report 2026*

## Traditionally, attackers have relied on four psychological tactics:

### Fear

Creating anxiety about potential consequences, such as losing access to an account or facing legal action.

### Urgency

Pressuring someone to act immediately without taking time to verify a request.

### Trust

Impersonating familiar people, companies, or authority figures to make a fraudulent request seem legitimate.

### Curiosity

Enticing someone to click a link, open an attachment, or investigate something unexpected.

**These tactics are still at the center of modern social engineering. What's changed is how effectively cybercriminals can use them.**



# Traditional vs. Modern Social Engineering

Before AI became widely accessible, many social engineering attacks required significant manual effort. Attackers had to research their targets, write convincing messages, and sometimes interact directly with victims to maintain their deception. **AI has made these activities faster, easier, and more scalable.**

| Traditional Social Engineering | Modern Social Engineering |
| --- | --- |
| Poorly written phishing emails | Polished, professional, localized messages |
| Generic mass phishing campaigns | Personalized attacks using publicly available information |
| Static phishing emails | Interactive, AI-powered conversations |
| Caller ID spoofing | AI-generated voice impersonation |
| Fake profiles and stolen photos | Realistic AI-generated images and videos |
| Manual research and impersonation | Automated research and content generation |

The result is a new generation of scams that can appear more credible and be deployed against more people with less effort. Traditional advice like checking for spelling mistakes or suspicious formatting is no longer enough. A fraudulent email can be perfectly written, a familiar voice can be artificially generated, and a video call with someone who looks exactly like your CEO may not actually involve your CEO.

Organizations need to **shift their focus** from teaching employees to recognize obvious mistakes **to helping them identify suspicious behaviors, requests, and situations.**

# Deepfakes: When Seeing and Hearing Are No Longer Believing

One of the most concerning developments in AI-powered social engineering is the rise of deepfakes.



**Deepfakes are synthetic or manipulated media created using AI to convincingly imitate real people, voices, or situations.**

**They generally fall into three categories:**

### Video

AI-generated or manipulated videos that make someone appear to say or do something they never actually said or did.

### Image

AI-generated or altered images that depict people or situations that may not be real.

### Audio

Synthetic audio that imitates a person's voice, accent, or speaking style.

While these technologies have legitimate applications, cybercriminals can use them to impersonate executives, manipulate employees, and create convincing fraudulent communications.

## Why Deepfakes Are So Effective

**People naturally rely on visual and auditory information to establish trust.** If we recognize someone's face or voice, we're more likely to believe we're communicating with that person. **Deepfakes exploit that instinct**, and research suggests that identifying synthetic content is far more difficult than many people realize.

# Can You Spot a Deepfake?

In a 2025 study of 2,000 UK and U.S. consumers...

**0.1%**

of participants correctly identified every real and AI-generated image and video shown to them.

**36%**

Participants were 36% less likely to correctly identify a deepfake video than a deepfake image.

*Source: iProov, 2025 Deepfake Detection Study*

Perhaps the most concerning part? **Participants were explicitly told to look for synthetic content before taking the test.** If people struggle to identify deepfakes when they're actively looking for them, imagine how difficult it becomes during an ordinary workday.

## THE TAKEAWAY:

Security awareness training shouldn't depend on employees becoming experts at recognizing AI-generated media. Instead, **organizations need verification procedures** that remain effective regardless of how realistic an impersonation becomes.

# Real-World Examples

## A $25 Million Deepfake Video Conference

**In 2024, engineering firm Arup became the victim of an elaborate deepfake scam involving approximately $25 million.**

The attack began when an employee in the company's Hong Kong office received a message that appeared to come from the UK-based chief financial officer. The message requested a confidential financial transaction.

Initially suspicious, the employee joined a video conference to verify the request. During the call, the employee saw what appeared to be the CFO and several colleagues, but the people on the call were AI-generated impersonations.

Convinced the request was legitimate, the employee transferred approximately $25 million across 15 transactions to five different bank accounts. The fraud was discovered only after the employee later contacted the company's head office.

*Source: Adaptive Security, Arup's $25M Deepfake Loss: Anatomy of an AI-Powered Scam*

## AI Voice Cloning Targets Italian Business Leaders

**In February 2025, several prominent Italian business leaders were targeted by scammers impersonating Italy's defense minister, Guido Crosetto.**

The attackers used AI-generated voice impersonation and calls that appeared to originate from government offices. Victims were told that Italian journalists had been kidnapped in the Middle East and that financial assistance was needed to secure their release. At least one target transferred approximately €1 million after believing the fraudulent request was legitimate.

The scam demonstrates how criminals can combine familiar voices, authority, and emotional manipulation to create convincing attacks.

*Source: The Guardian, February 10, 2025*

**A familiar voice, recognizable phone number, or seemingly legitimate authority figure should never replace independent verification, and even employees who attempt to verify suspicious requests can be deceived when the verification process itself is compromised. Organizations need independent verification channels and established approval procedures for sensitive transactions.**

# Business Email Compromise: An Old Threat Gets More Convincing

Not every successful social engineering attack requires deepfake technology. Sometimes, all it takes is a believable email.

**Business email compromise (BEC) is a form of fraud in which attackers use fake or hacked email accounts to impersonate trusted individuals or organizations** and trick someone into transferring money, revealing sensitive information, or granting unauthorized access.

Common targets include executives, finance teams, HR departments, and employees responsible for vendor payments. BEC has been a major cybersecurity threat for years, but AI is making it easier for criminals to create more convincing impersonations.



## How AI Is Changing Business Email Compromise

Traditional BEC attacks often involved manually researching a target and crafting a message that appeared to come from a trusted person. AI can accelerate this process by helping attackers:

### Create convincing messages

Generate professional emails that match the tone and language of legitimate business communications.

### Personalize attacks

Incorporate publicly available information about employees, organizations, and business relationships.

### Scale campaigns

Produce customized messages for multiple targets without manually writing each one.

### Adapt communications

Generate responses that maintain a believable conversation when recipients ask questions.

**The result is a type of fraud that may look nearly identical to an ordinary business transaction.**

Consider a company receiving an email from a familiar vendor:

> *"We've recently updated our banking information. Please use the attached payment details for your next invoice."*

**There's no obvious threat, suspicious grammar, or urgent demand.**

But if the message came from an attacker impersonating the vendor, the company could unknowingly redirect a legitimate payment into a fraudulent account, and the financial consequences could be significant.

## The Cost of Business Email Compromise

**24,768**

BEC complaints reported to the FBI's Internet Crime Complaint Center in 2025

**$3.047B**

Reported BEC losses in 2025

**#2**

BEC was second only to investment fraud among cyber-enabled fraud categories by reported financial losses

*Source: FBI, 2025 Internet Crime Report*

# Real-World Examples

## $7.5 Million Stolen from Charity

**In a case involving two U.S. charitable organizations, an attacker allegedly gained unauthorized access to email accounts and impersonated employees of one charity to initiate fraudulent financial transactions with another charity.**

The attacker used compromised credentials, spoofed domains, and manipulated email communications to make unauthorized transfers appear legitimate.

The scheme resulted in approximately $7.5 million being transferred to accounts controlled by the attacker.

The case illustrates how criminals can exploit existing business relationships and trusted communication channels to bypass financial safeguards.

*Source: BleepingComputer, January 3, 2024*

## North Dakota School District Loses $4.9 Million

**In February 2026, Dickinson Public Schools in North Dakota disclosed that it had lost approximately $4.92 million to an email scam.**

"Criminals impersonated a trusted vendor and redirected two payments from the district's building fund [into a fraudulent account]."

The incident demonstrates how even routine vendor relationships can become opportunities for financial fraud.

*Source: Government Technology, February 13, 2026*



**Email authentication and account security are essential, but organizations also need independent approval processes for financial transactions. Changes to vendor payment information should always trigger additional verification, regardless of how familiar or legitimate the communication appears.**

# Five Red Flags of Modern Social Engineering

With the rapid evolution of social engineering scams, you might expect the warning signs to have evolved too. Surprisingly, they haven't.

**The same psychological tactics that made traditional social engineering successful are still being used today.** For MSPs, these five red flags should remain central to security awareness training.

### 1 Urgency

Attackers create a false sense of urgency to discourage people from thinking critically.

*"Your account will be suspended unless you verify your information immediately."*

### 2 Pressure

Scammers may discourage employees from asking questions, following procedures, or involving others.

*"This is confidential. Please handle it directly and don't involve anyone else."*

### 3 Threats

Attackers may use fear of consequences to manipulate decisions.

*"Failure to complete this could result in legal action being taken."*

### 4 Financial Requests

Unexpected requests involving money, payment changes, or financial information deserve additional scrutiny.

*"We've changed banks. Please update our payment information before processing the next invoice."*

### 5 Emotional Manipulation

Scammers exploit fear, sympathy, excitement, and personal relationships to influence behavior.

*"I've been in an accident and need help. Please send money as soon as you can."*

**AI simply makes these warning signs easier to disguise.**

While these warning signs remain relevant, organizations can't assume every fraudulent communication will contain an obvious red flag. Some of the most convincing attacks resemble perfectly ordinary business requests.

**That's why effective security awareness training should teach employees to evaluate the context of a request, not just the appearance of a message.**

## Employees should always be asking themselves:

- **Was the request expected?**
- **Does it follow normal procedures?**
- **Is someone asking for an unusual action?**
- **Can the person's identity be verified independently?**

**These questions are increasingly important in a world where appearances can be manufactured.**

# The Business Risks of AI-Powered Social Engineering

Social engineering is more than an employee awareness issue. A successful attack can create **financial, operational, and reputational consequences** that extend across an organization. For MSPs, understanding these risks is essential when helping clients evaluate their cybersecurity investments.

## 01 Financial Loss

BEC attacks can redirect legitimate payments, while phishing and impersonation attacks can lead to unauthorized transactions, stolen credentials, and costly data breaches.

**Financial loss is one of the most immediate consequences of successful social engineering.**

Beyond the initial financial loss, organizations may face additional expenses related to investigations, recovery, legal assistance, and security improvements.

## The Financial Cost of Breaches

**$4.99M**

Global average cost of a data breach in 2026

**$6M**

Average cost of an AI-enabled malicious breach

*Source: IBM, Cost of a Data Breach Report 2026*

## 02 Reputational Damage

When an organization experiences a security incident, the consequences can extend beyond its financial statements. Customers may question whether their personal information is secure, business partners may become hesitant to share sensitive data, and employees may lose confidence in the organization's ability to protect its systems.

**Rebuilding that trust can take significantly longer than resolving the initial incident.**

For MSPs, this is an important reminder that cybersecurity is closely tied to customer relationships and business continuity.

## 03 Business Disruption

A successful social engineering attack can also interrupt everyday operations. Compromised accounts may need to be disabled, systems investigated, and employees temporarily restricted from accessing critical resources.

Finance teams may need to review transactions, IT teams may spend days investigating suspicious activity, and leadership may be pulled away from normal business priorities.

## The Time Required to Recover

**247 days**

Average time to identify and contain a data breach in 2026

*Source: IBM, Cost of a Data Breach Report 2026*

**The longer an incident goes undetected, the greater the potential for additional damage and operational disruption.**

For MSPs managing multiple client environments, incidents can be even more detrimental to your workload.

# How MSPs Can Build Stronger Defenses Against AI-Powered Social Engineering

The most effective defense against modern social engineering isn't necessarily a completely new security strategy. It's getting back to the basics: **strengthening the fundamentals that have always mattered and using modern technology to make them more effective.**

## Here are the four security layers every MSP should consider.

**LAYER 1**
**Multi-Channel Verification**

**LAYER 2**
**Continuous Security Awareness Training**

**LAYER 3**
**Established Verification Procedures**

**LAYER 4**
**AI-Powered Detection and Security Tools**

## Layer 1: Multi-Channel Verification

**One of the strongest defenses against impersonation is independent verification.**

If an employee receives an unusual request involving money, credentials, or sensitive information, they should confirm the request through a separate, trusted communication channel.

For example, if a vendor emails updated banking information, the employee should contact the vendor using a phone number already on file. If an executive requests an urgent financial transaction, the employee should verify it through an established internal process.

The important distinction is that verification must happen independently of the suspicious communication. A video call initiated by an attacker isn't independent verification, even if the participants look familiar.

**Help clients establish clear verification expectations for high-risk communications, particularly those involving financial transactions, account access, and sensitive information.**

## Layer 2: Continuous Security Awareness Training

**Training should focus on helping employees recognize suspicious behaviors, question unusual requests, and respond appropriately.**

That includes understanding how criminals use urgency, trust, fear, and emotional manipulation to influence decisions. Training should also cover emerging tactics like deepfakes, voice cloning, and AI-generated phishing emails.

But awareness shouldn't be limited to an annual training session. Regular, short training sessions and phishing simulations can help reinforce safer habits over time, and when someone falls for a simulated phishing email, the goal should be education rather than punishment.

**Build a continuous training program that reinforces practical security behaviors and measures improvement over time, rather than simply tracking course completion.**

## Layer 3: Established Security Procedures

Awareness is important, but employees shouldn't be expected to make every security decision on their own. **Organizations need clear procedures that make it easier to identify and respond to unusual requests.**

**Work with clients to identify high-risk workflows and implement verification procedures that employees can realistically follow. Even something as simple as implementing a codeword can make a huge difference.**

For example:

- Require additional approval for large financial transactions.
- Independently verify changes to vendor payment information.
- Establish procedures for handling unexpected requests from executives.
- Use multi-factor authentication to protect sensitive accounts.
- Create clear processes for reporting suspicious communications.

These procedures help reduce reliance on individual judgment.

# Layer 4: AI-Powered Detection and Security Tools

**AI might be making social engineering attacks more sophisticated, but it can also help organizations defend against them.**

Security teams can use AI and automation to identify suspicious activity, analyze potential threats, and reduce the manual work involved in investigating incidents.

For MSPs, this can be particularly valuable when managing security across multiple client environments.

Some important capabilities include:



### Email authentication and protection

Technologies such as SPF, DKIM, and DMARC help organizations authenticate email and reduce certain types of domain spoofing. Email security tools can provide additional protection by blocking malicious content before someone ever has the chance to make a mistake.



### Automated phishing analysis

AI-powered tools can help analyze suspicious emails, examine headers and URLs, assess potential threats, and provide technicians with information needed to investigate reported messages. Automating phishing analysis and prioritization can reduce technician fatigue and help teams focus on higher-risk threats.



### Dark web monitoring

Monitoring for exposed credentials can help MSPs identify potential account risks and take action before stolen information is used in an attack.



### Automated workflows

Integrating security tools with existing service desk and ticketing systems can help reduce repetitive tasks and accelerate incident response.

**These capabilities don't eliminate the need for human oversight.** Instead, they help security teams work more efficiently and focus their attention where it's needed most. Research suggests the impact can be significant.

# The Advantage of AI and Automation

## 65 days

**FASTER**

Organizations extensively using AI and automation identified and contained breaches 65 days faster than organizations that didn't use these technologies.

## $1.93M

**LESS IN BREACH COSTS**

Average breach costs for organizations extensively using AI and automation compared with those using none.

*Source: IBM, Cost of a Data Breach Report 2026*

**Evaluate where AI and automation can reduce manual security work, accelerate investigations, and improve visibility across client environments.**

### These findings demonstrate that AI isn't exclusively a tool for attackers.

When implemented appropriately, it can also strengthen security operations and help organizations respond to threats more efficiently.

# Fighting AI-Powered Threats Requires More Than AI-Powered Tools

Social engineering has entered a new era. **AI has made it easier for criminals** to create believable messages, impersonate trusted individuals, and launch sophisticated attacks at scale, **but the underlying psychology remains the same.**

**Attackers still exploit trust, urgency, fear, and curiosity** to manipulate people into making decisions they otherwise wouldn't. The opportunity for MSPs isn't simply to introduce more security tools or tell employees to be more careful. It's to **build a security strategy that combines people, processes, and technology.**



**Help employees recognize suspicious situations, establish procedures that reduce opportunities for fraud, and implement tools that make threats easier to detect and investigate.**

Employees shouldn't be expected to catch everything. As impersonation technology becomes more convincing, organizations need security layers that protect people even when they can't distinguish a legitimate communication from a fraudulent one.

**AI has changed the threat landscape. It's time to make sure your security strategy keeps up.**

# Build Stronger Security Layers with Phin

**AI is making cyber threats more sophisticated, but it can also make your security operations more efficient.**

Phin helps MSPs strengthen their clients' defenses with continuous security awareness training, automated phishing simulations, AI-powered phishing analysis, and dark web monitoring.

With **Phinbox IQ**, MSPs can use AI to analyze reported emails, reduce manual phishing triage, and help technicians investigate potential threats faster. Combined with ongoing training and automated workflows, Phin helps MSPs build stronger security habits while reducing the time spent managing them.

---

SAVED

## 50+ hrs/week

One MSP saves 50+ hours of tech time per week with Phin's automated tools, increasing profitability and security with faster resolutions. Many other MSPs are experiencing similar results.

Source. How an MSP Saved 50+ Hours Per Week With Phinbox IQ

*"Phin's reduced the human error rate... not just on the end user side, but also on the engineer side."*

- Joel Chambers, Escalation Specialist at Certified CIO

## Start putting AI and automation to work for your MSP.

**Explore Phin's Products →**

PHIN