

State of the Human Attack Surface

2026 Threat Intelligence Report

Summary

01

Executive Summary

02

The End of Trust as a Security Control

03

The Anatomy of a Modern Social Engineering Attack

04

The Orchestration Problem

05

Implications for Security Leaders

06

About Adaptive Security

Guides, videos, blogs, and product updates to help you build a security-aware workforce.

[WWW.ADAPTIVESECURITY.COM/RESOURCES](https://www.adaptivesecurity.com/resources)

Executive Summary

The human attack surface is the primary battleground of modern enterprise security. AI has given attackers a capability set that enterprises were never built to defend against. Technical controls alone cannot close that gap.

Three forces are reshaping the 2026 social engineering threat:

- AI has industrialized attack creation. Campaigns that once required hours of manual reconnaissance now launch in minutes, personalized, convincing, and deployed at scale.
- Attack vectors no longer operate in isolation. Threat actors now chain email, SMS, voice, and video into coordinated kill chains designed to outlast any single-channel defense.
- The identity layer has collapsed. Voice cloning, synthetic video, and AI-generated impersonation have made the verification signals employees once relied on actively exploitable.

Recent cybersecurity statistics and data breach statistics confirm that social engineering cyberattacks remain the primary catalyst for corporate network intrusions, driving the average cost of a data breach to record levels. The Mandiant M-Trends 2026 Report documents that voice-centric deception tactics now constitute 11% of analyzed business network intrusions, cementing vishing's position as the runner-up mechanism for global threat actor entry. This pairs with data from IBM showing that roughly one in six data breaches now involves threat actors actively deploying AI tools, primarily across phishing, deepfake impersonation, and identity-focused attacks.

These statistics confirm what security teams are already experiencing: Phishing trends in 2026 reflect a threat landscape that has fundamentally shifted, with individual attack vectors converging into coordinated, multi-channel kill chains that exploit the seams between enterprise defenses.

This human attack surface report examines how AI social engineering exploits the seams between enterprise defenses across eight critical vectors: AI-powered phishing, QR code phishing (quishing), SMS phishing (smishing), deepfake AI impersonations, voice phishing (vishing), help desk deception, collaboration platform compromise, and business email compromise (BEC).

To counter these cross-channel threats, organizations must transition from siloed technical perimeters to an integrated human-layer defense strategy. Adaptive Security's platform bridges this structural gap by unifying multi-channel phishing simulations, automated triage, and behavioral risk monitoring into a single framework for enterprise resilience.

The End of Trust as a Security Control

A high-priority email landed in the inbox of an enterprise finance director: 'We need to finalize the regional acquisition funding immediately. Let's jump on a video call to authorize the \$25 million transfer.' The sender profile perfectly matched the company's global CFO.

When the video conference launched, any initial hesitation vanished. The CFO's face, vocal cadence, and physical gestures were flawless. Two other internal colleagues were visible on screen with him. The context was entirely plausible, and the instructions were clear. Within an hour, the director authorized the multi-million-dollar wire, completely unaware that the entire meeting was an orchestrated attack utilizing real-time, AI-generated deepfake clones of the executive team.

This isn't a hypothetical. In 2024, engineering firm Arup lost \$25 million to exactly this attack. The deepfakes were convincing enough that a finance employee authorized the transfer without hesitation. Enterprises have spent a decade hardening infrastructure via zero-trust frameworks, endpoint detection and response, and secure email gateways. Attackers have simply moved around it, targeting the people inside instead.

2026 VERIZON DATA BREACH INVESTIGATIONS REPORT

The human element continues to be a factor in 60% of all confirmed breaches.

Generative AI has fundamentally compressed the threat timeline:

Attackers can clone an executive's voice from seconds of public audio and launch hyper-personalized phishing campaigns in five minutes instead of sixteen hours. This acceleration aligns with broader network infrastructure trends tracked in the Mandiant M-Trends 2026 Report, which documented a global median attacker dwell time of 14 days.

[LEARN MORE](#)

That window is shrinking. Security teams must now intercept human-layer exploits before adversaries establish internal persistence. Modern adversaries no longer rely on isolated vectors. They chain email, SMS, voice, video, and collaboration platforms into a single coordinated sequence, designed to find the one moment an employee's guard is down. The human layer is the most targeted and least-defended asset in the enterprise.

The Anatomy of a Modern Social Engineering Attack

Modern social engineering attacks follow a three-phase kill chain: Initial access, identity exploitation, and payload execution. Every phase targets a different layer of enterprise trust, and none succeeds without the one before it.

PHASE 1

Initial Access

Attackers establish their first foothold through three scalable entry points: email, SMS, and QR code.

1. AI-Powered Phishing

Phishing remains the most prevalent initial access vector in corporate breaches, but the underlying methodology has shifted. Generative AI has converted generic, mass-volume spam into high-precision spear phishing at industrial scale. Adversaries now pair automated reconnaissance with compromised organizational data to generate lures that replicate an enterprise's internal communication style with near-perfect fidelity.

This tactical evolution is clear in the [Storm-2372 device code phishing campaign](#). Rather than deploying traditional malicious hyperlinks, threat actors used highly convincing Microsoft authentication device-code prompts targeting global enterprises, NGOs, and government networks. Because this payload requires zero malicious link clicks, it bypassed secure email gateways and user suspicion.

Legacy security awareness training built on detecting typos or checking URLs is obsolete against AI-generated attacks engineered to exploit legitimate corporate authentication workflows.

2. QR Code Phishing (Quishing)

QR code phishing has become one of the fastest-growing initial access vectors in corporate environments.



According to the [Microsoft Threat Intelligence Q1 2026 Trend Report](#), malicious QR code attacks spiked 146% in early 2026, reaching over 18 million detections monthly.

Quishing works because it was designed to evade the traditional security stack. Attackers embed malicious destinations inside QR images delivered via email or PDF attachments, a method the Anti-Phishing Working Group (APWG) Trends Report documents has scaled aggressively across the enterprise sector. Because the payload is concealed within a graphic asset and resolved on an employee's personal mobile device, secure email gateways and URL inspection tools have nothing to parse.

The attack itself is straightforward. An employee receives what appears to be a routine corporate re-authentication request. They scan the code on their phone, get redirected to a convincing counterfeit portal, and hand over their credentials and session tokens on a device no IT team is watching.

3. SMS Attacks (Smishing)

Mobile devices are the least monitored environment in the enterprise.

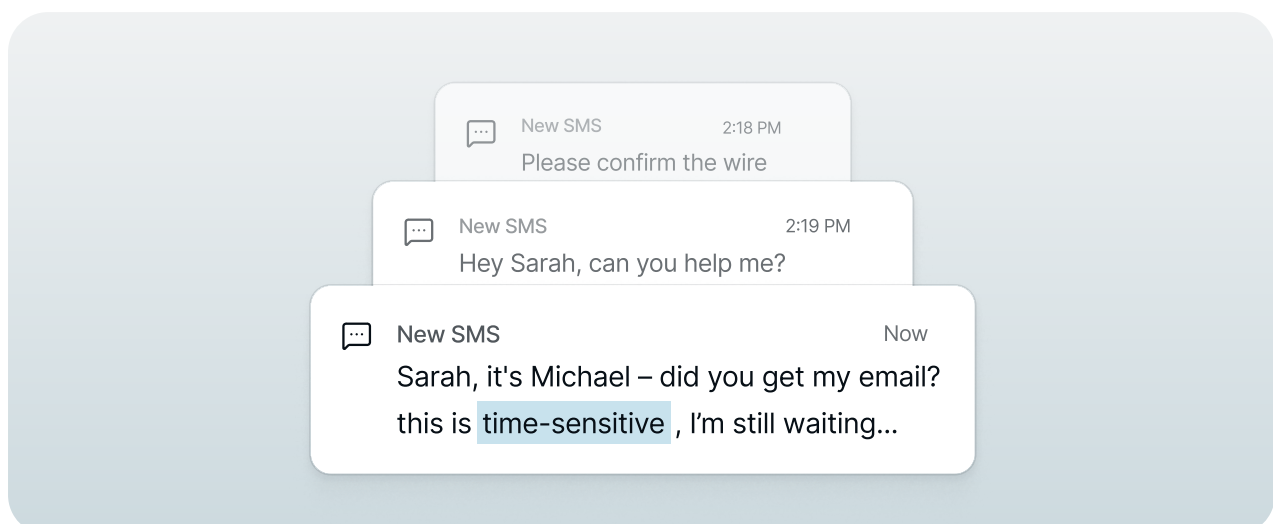
~70%

The Verizon Mobile Security Index (MSI) reports that ~70% of IT administrators acknowledge at least half the devices accessing corporate networks are unmanaged.

~50%

The 2026 Verizon DBIR puts a finer point on the exposure: ~50% of compromised corporate credentials originated from unmanaged BYOD endpoints.

Smishing exploits that gap directly. SMS arrives on personal screens without corporate carrier filtering, and it works. The same Verizon Report found that phone-channel simulations yield click-through rates nearly double those of corporate email. There is no enterprise control standing between an attacker's message and an employee's thumb.



PHASE 2

Identity Exploitation

Following initial access, attackers use generative AI to construct synthetic voice and video replicas of trusted internal identities. The result is impersonation at a fidelity that traditional verification controls were never designed to catch.

1. Deepfake Attacks

Deepfakes have become a production-ready enterprise weapon. Attackers deploy voice cloning for executive impersonation and synthetic video to bypass identity verification. Modern deepfake tools require as little as three seconds of source audio to produce a replica convincing enough to authorize a fraudulent wire transfer.



A Gartner User Survey found that 62% of organizations have experienced a deepfake attack targeting automated workflows or internal teams within a 12-month window. Gartner Research separately projects that 30% of enterprises will consider biometric identity verification completely unreliable when deployed in isolation, as standard liveness checks fail to parse high-quality synthetic injection attacks.

To counter this, organizations are moving beyond legacy filters and deploying AI and deepfake threat training to build behavioral resilience against advanced synthetic identity fraud.



8 photos, 2 video clips ✓



3 voice clips ✓



2. Voice Phishing (Vishing)

11%

The [Mandiant M-Trends 2026 Report](#) revealed that interactive voice phishing jumped to 11% of all investigated corporate intrusions, making vishing the second-most common initial infection vector globally.

73%

This trajectory aligns with the [World Economic Forum \(WEF\) Global Cybersecurity Outlook 2026](#), which notes that 73% of organizations have now actively detected vishing, smishing, or advanced phishing variants within their enterprise perimeters.

Generative AI has transformed voice phishing from a low-scale manual operation into an industrialized one. Attackers scrape publicly available executive audio to engineer live voice clones that match speech patterns, accent, and emotional inflection. They use those clones to pressure employees into bypassing verification safeguards at high-stakes moments: MFA resets, payroll updates, urgent wire transfers.

Employees have spent years learning to spot suspicious emails. Almost none have been trained to challenge a live call from someone who sounds exactly like their CEO.

3. Help Desk Impersonation

Designed to resolve urgent access issues rapidly, help desks are primary targets for attackers who recognize that manipulating identity verification workflows is easier than attacking technical infrastructure directly.

The attack methodology relies on precise reconnaissance: Attackers aggregate employee data from open-source intelligence, credential dumps, and corporate directories. Armed with legitimate contextual details (e.g., job titles, active projects, or partial credentials), they contact help desk personnel while impersonating high-value internal staff. Their objective is to force unauthorized password resets, malicious multi-factor authentication (MFA) re-enrollments, or fraudulent device registrations.



The [MGM Resorts breach](#) is the defining case:

attackers manipulated the help desk into granting elevated administrative permissions, costing the organization over \$100 million. [Microsoft](#) has since tracked persistent threat groups doing the same thing systematically. Calling the help desk gets attackers further than attacking the firewall.

PHASE 3

Payload Execution (Financial and Operational Manipulation)

With credibility established through synthetic identity exploitation, attackers move to collect. The final phase converts access into money, data, or operational disruption.

1. Collaboration Platform Impersonation

Employees treat Microsoft Teams, Slack, and Zoom as trusted, authenticated environments. That trust is the vulnerability. The informal, real-time cadence of these platforms lowers skepticism in ways that a suspicious external email never would. A message appearing to come from internal IT or a senior executive triggers compliance, not scrutiny.

According to [documented Microsoft technical advisories](#), advanced persistent threat groups (including the Russian state-sponsored espionage group Midnight Blizzard and the financially motivated cybercrime collective Scattered Spider) frequently exploit external tenant messaging features to blur organizational boundaries. In these campaigns, attackers launch tenant-spoofing Teams messages and malicious follow-up support calls to pose as internal help desk personnel. By convincing employees to grant remote device access via native tools like Quick Assist, attackers easily bypass legacy perimeter defenses, deploy ransomware payloads, and establish persistent footholds.



Business email compromise (BEC) ranks as the [second most financially destructive cybercrime](#) in the United States. The [FBI IC3 2025 Annual Report](#) documented over \$3 billion in verified BEC losses across approximately 24,000 corporate complaints, pushing cumulative global losses past \$55 billion over the last decade.



Attackers succeed by manipulating high-privilege employees in financial or operational roles. Approximately 40% of [BEC phishing emails](#) are now AI-generated, optimized for tone consistency and writing style mimicry to bypass traditional human scrutiny.



The [2026 Dickinson Public Schools case](#), which saw \$4.85 million in fraudulent transfers initiated before federal law enforcement intervened, demonstrates the massive financial exposure facing non-banking sectors. According to the Verizon 2026 DBIR, pretexting has nearly doubled in frequency, positioning credential-driven fraud as a primary cause of enterprise breaches.

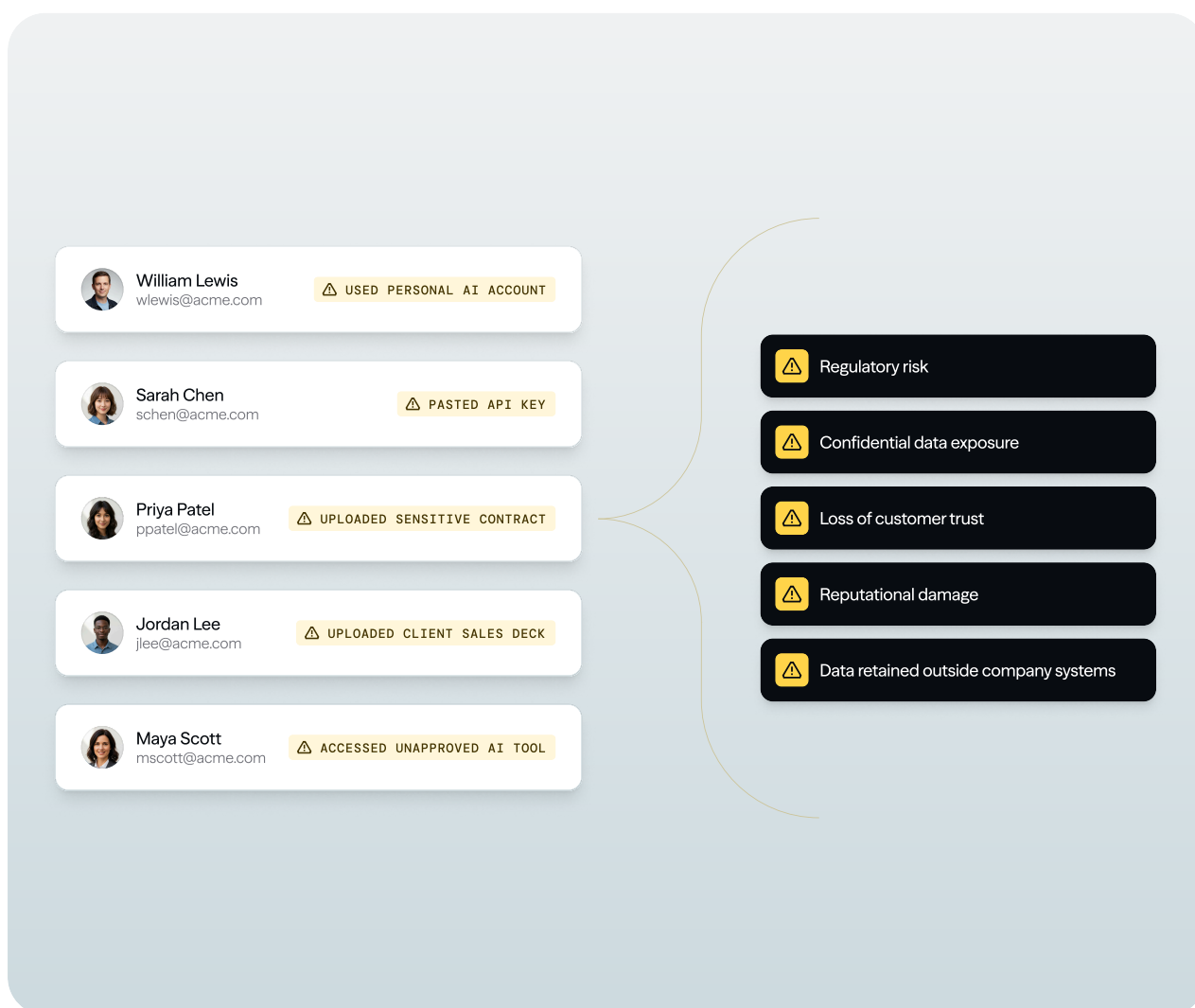
The Orchestration Problem

The three phases of a modern social engineering attack don't operate independently. They are designed to connect, with each one building the credibility the next one needs to succeed.

This is what that looks like in practice: an AI-generated invoice arrives via email, an SMS creates urgency, a cloned voice call confirms the request, and a Teams message from a spoofed internal account closes it. No single channel carries the full attack. The kill chain lives in the seams between them.

Attackers succeed because they move across channels faster than disconnected technical controls can correlate the activity, and faster than any employee can manually verify what's happening. Security architectures that treat email filtering, mobile security, and identity verification as separate problems are blind to campaigns specifically engineered to exploit that separation.

Single-channel defenses were built for a different threat. The attack has moved on.



Implications for Security Leaders

The FBI introduced a dedicated AI-enabled crime category in its 2025 IC3 report for the first time, logging 22,364 complaints and \$893 million in confirmed losses. That figure almost certainly understates actual exposure, as most deepfake fraud goes undetected or unreported. Generative AI has given attackers the tools to move faster, imitate more convincingly, and operate across more channels than legacy security models were designed to handle.

When AI-driven impersonation attacks succeed by replicating a recognized CFO's voice, mimicking a trusted vendor's writing style in an active invoice thread, or appearing as a colleague inside familiar collaboration platforms, employees require far more than recycled awareness training or basic tooling upgrades to spot malicious actors.

The following recommendations provide a practical starting point for leaders looking to close the human-layer gap against AI-enabled threats:



Unify your defenses across every channel.

Multi-channel attacks cannot be countered by channel-specific controls operating independently. Centralized telemetry, cross-channel simulation, unified reporting, and stronger authentication requirements for unmanaged BYOD devices are baseline requirements for organizations facing coordinated kill chains. The architecture has to match the threat.



Expand and diversify simulation training.

AI-generated social engineering unfolds across SMS, voice, QR codes, collaboration platforms, and video, often within the same attack sequence. Training calibrated only for suspicious email links leaves employees exposed across every other channel. Simulations must reflect how attacks actually arrive, including dedicated scenarios for vishing and smishing, which see click-through rates nearly double those of email.



Supplement MFA and identity verification with out-of-band protocols.

Deepfakes and AI-assisted vishing have invalidated biometric and voice-based identity verification when used in isolation. High-value actions (e.g., funds transfers, payroll changes, vendor modifications, and admin credential resets) must require step-up authentication, callback verification, or hardware-backed verification that cannot be replicated through synthetic impersonation.

**Apply risk-based escalation to high-value workflows.**

Finance, procurement, payroll, and admin credential workflows are disproportionately targeted. CISOs must establish escalation controls that require secondary out-of-band confirmation for transactions above defined thresholds. This will reduce the window for social engineering to succeed even after initial trust is established.

**Adopt a unified human-layer defense strategy.**

Multi-channel attack patterns cannot be countered by channel-specific controls operating independently. Centralized telemetry, cross-channel phishing simulation, unified reporting, and stronger authentication requirements for unmanaged BYOD devices are baseline requirements for organizations facing coordinated kill chains. The speed of AI-enabled attacks demands that human recognition and verification capabilities improve materially.

**Establish human-layer metrics.**

Most organizations measure technical controls rigorously and human-layer controls almost not at all. CISOs should track simulation click and report rates by channel, employee-initiated reporting rates as a share of total simulated attacks, mean time to report across departments and seniority levels, and escalation adherence rates on high-value workflows. These indicators reframe human-layer defense as an improving capability rather than a static one.

Organizations that close the human-layer gap will not be those that deploy the most tools, but those that treat employee recognition, training, and verification as a measurable capability: one that is trained, tested, and iterated on with the same rigor applied to technical controls.

Edit Test Out

☰ Test Out ▼

☰ Content ^

Introduction

Integrity in Code

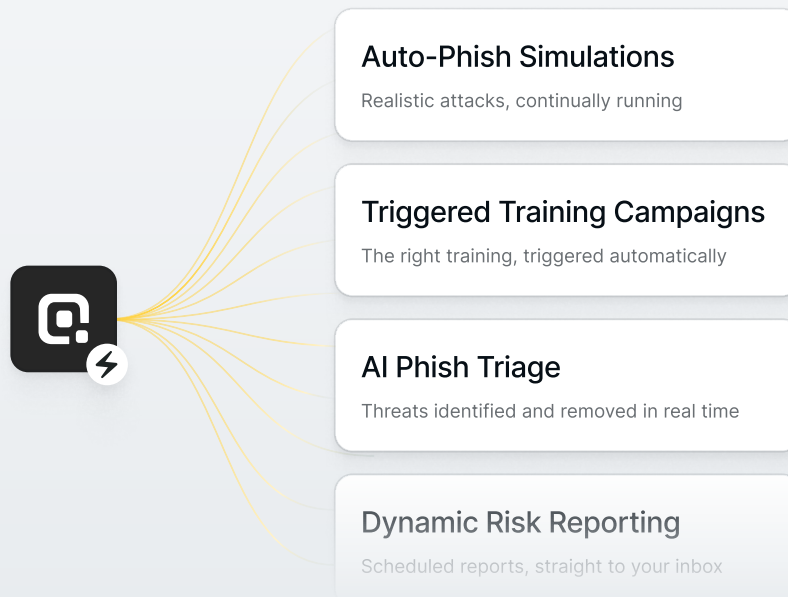
Scenarios & Specifics

Prevention Measures

Quiz

Question 2☐ They use older, less secure devices than tenured employees☒ They haven't yet learned the company's security protocols☐ They have more system permissions during onboarding

Submit and review results



About Adaptive Security

Adaptive Security is the security layer for the AI era. The attacks documented in this report are not theoretical. Deepfakes, AI-generated spear phishing, vishing, and coordinated multi-channel kill chains are the threats organizations are facing today, and they are outpacing the defenses most enterprises have in place.

Adaptive addresses the full scope of that threat in one unified platform: training employees on the attacks they will actually face, simulating AI-powered threats across every channel, protecting the inbox from what native filters miss, and giving security teams visibility into the AI tools their employees are already using.

Because it is one platform, every product shares data and makes the others smarter. An email near-miss triggers targeted training before the next attack lands. Simulation results shape which employees get prioritized for intervention. The result is a human risk program that gets more accurate and more automated over time, closing the gaps that multi-channel attackers are designed to exploit.

More than 1,000 organizations run Adaptive to manage their human security program. The company has raised \$150M+ from investors including NVIDIA, Andreessen Horowitz, Bain Capital Ventures, and Citi Ventures.

Take the next step in security awareness training and prepare your teams for the future of AI cybersecurity.

TO LEARN MORE, VISIT [WWW.ADAPTIVESECURITY.COM](https://www.adaptivesecurity.com)