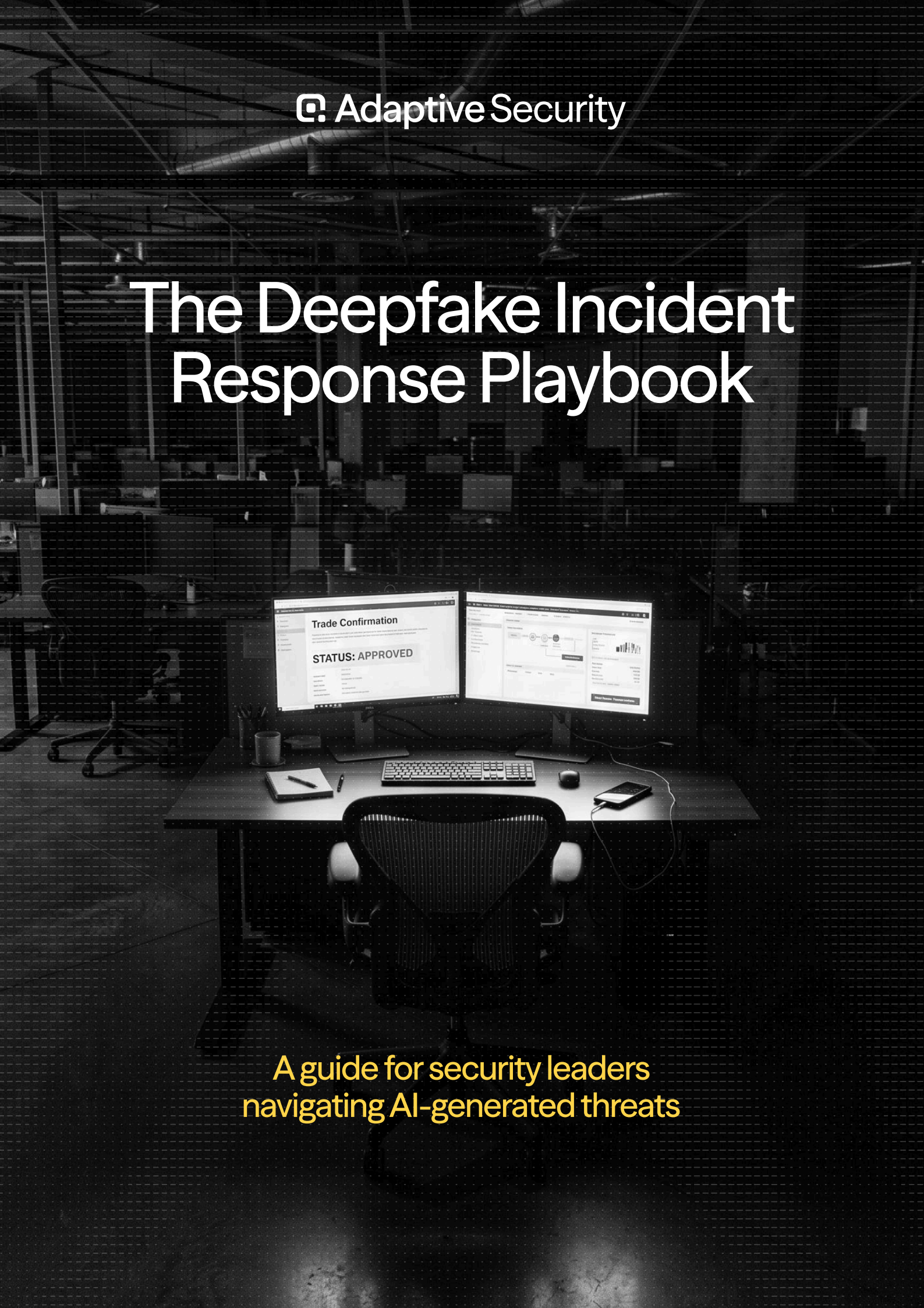
Adaptive Security

The Deepfake Incident Response Playbook



A guide for security leaders navigating AI-generated threats



In This Playbook

In July 2025, an impersonator used an AI-generated voice clone of Secretary of State Marco Rubio to contact foreign ministers, a U.S. governor, and a member of Congress through Signal. The State Department opened an investigation. The voicemails were convincing enough to require official government-wide advisories to foreign partners. No sensitive information was extracted in that instance, but the same technique has been deployed successfully in financial operations.

In March 2025, a finance director at a multinational firm in Singapore joined what appeared to be a routine Zoom call with company leadership. The CFO and several senior executives were on screen. Their voices, faces, and mannerisms matched. At the end of the call, the finance director authorized a $499,000 transfer. Every person on that call was a deepfake.

Both cases share a common thread. Employees were asked to make a trust judgment they had no framework for. UC Berkeley digital forensics professor Hany Farid has documented why: untrained people perform only marginally better than chance when asked to identify AI-generated content. His research predates the era of real-time, interactive deepfakes now widely accessible to bad actors.

This playbook is built for the security teams responsible for closing that gap. It covers the five deepfake attack types organizations encounter most frequently, what each looks like in practice, and what a well-executed response looks like from initial detection through post-incident review. The framework in Section 2 applies across all five scenarios. The recognition signals and response steps are specific to each.

Security teams that build detection skills before an incident, establish clear verification protocols, and rehearse cross-functional response contain these attacks faster and recover with less damage when one gets through. That is the purpose of this guide.

What this playbook covers:

- **How AI-generated threats operate across five distinct attack surfaces**
- **A four-phase incident response framework applicable across all deepfake scenarios**
- **Step-by-step response guidance for each incident type, phased by timeline**
- **A proactive defense program built around training, verification, and tabletop exercises**
- **Cross-functional response team roles and post-incident review practices**
- **How Adaptive Security supports each phase of your program**

Summary

01

The Deepfake Attack Surface

02

Your Response Framework

03

Five Incident Types: Detection to Resolution

04

Building a Proactive Defense Program

05

Your Response Team and Post-Incident Review

06

How Adaptive Security Partners With Your Team

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

ADAPTIVE RESOURCES

SECTION 1

The Deepfake Attack Surface

Deepfake attacks increased 17x in the United States from 2023 to 2024, and the trend hasn't slowed. Security professionals reported more than 100,000 incidents in the U.S. in 2024 alone, and that figure understates the problem: most successful enterprise attacks go undisclosed. The FBI's 2025 Internet Crime Report ties at least $893 million in losses to AI-enabled fraud last year — including more than $5 million from cloned-voice scams impersonating a family member in distress — and calls that figure a floor, since many AI-enabled cases are never flagged as such at intake. Fraud, impersonation, and extortion now drive most of what this technology gets used for, and CISOs are managing a volume and sophistication of AI-powered threats that public reporting does not fully capture.

The technology behind those attacks has become both cheap and widely accessible. Voice cloning now requires only a few minutes of source audio. Deepfake video tools require no technical background. The barrier to entry dropped so sharply between 2023 and 2024 that, as Adaptive Security CEO Brian Long has observed, the capability is available to virtually anyone. Intent is now the only limiting factor.

What Changed, and What Stayed the Same

Deepfake attacks exploit the same psychological levers as every other social engineering attempt: authority, urgency, and familiarity.

Ethical hacker Rachel Tobac, who has demonstrated these techniques live on CNN and 60 Minutes, puts it plainly: attackers combining authority and urgency can get someone to act on a malicious request in under 30 seconds.

An employee who receives a video message from their CEO requesting an urgent wire transfer faces the same pressure as a well-crafted phishing email, with one added element. They can see and hear a trusted face and voice. The delivery mechanism is new. The underlying manipulation is decades old.

That distinction matters because your team already has the foundation to respond. Verification protocols, clear escalation paths, and employees trained to slow down under pressure are effective against deepfakes for the same reason they work against any social engineering attempt. What requires specific preparation is learning to recognize the signals that distinguish a deepfake from a legitimate interaction, understanding which roles carry the most exposure, and building practiced response steps for the scenarios your organization is most likely to face.

Four Ways Deepfakes Enter an Organization

Impersonation	Cloning an executive's voice or video likeness to authorize transactions, request credential resets, or push employees toward actions outside normal approval workflows.
Manipulation	Using AI-generated documents, fabricated context, or synthetic personas to reinforce a broader social engineering campaign.
Infiltration	Creating AI-generated identities to pass hiring processes and gain internal access. Observed in nation-state operations and financially motivated fraud.
Disinformation	Generating fabricated content attributed to executives or employees to erode internal trust or damage an organization's external reputation.

These four categories frame the five specific scenarios in Section 3. The response framework that follows applies across all of them.

SECTION 2

Your Response Framework

Every deepfake incident, regardless of type, moves through four phases. How quickly and cleanly a team moves through each phase determines whether an incident stays contained or expands into something larger.

1. Recognize

Identify signals suggesting a deepfake may be involved. Engage the right people immediately. Preserve all available evidence before anything is deleted, expired, or overwritten.

2. Contain

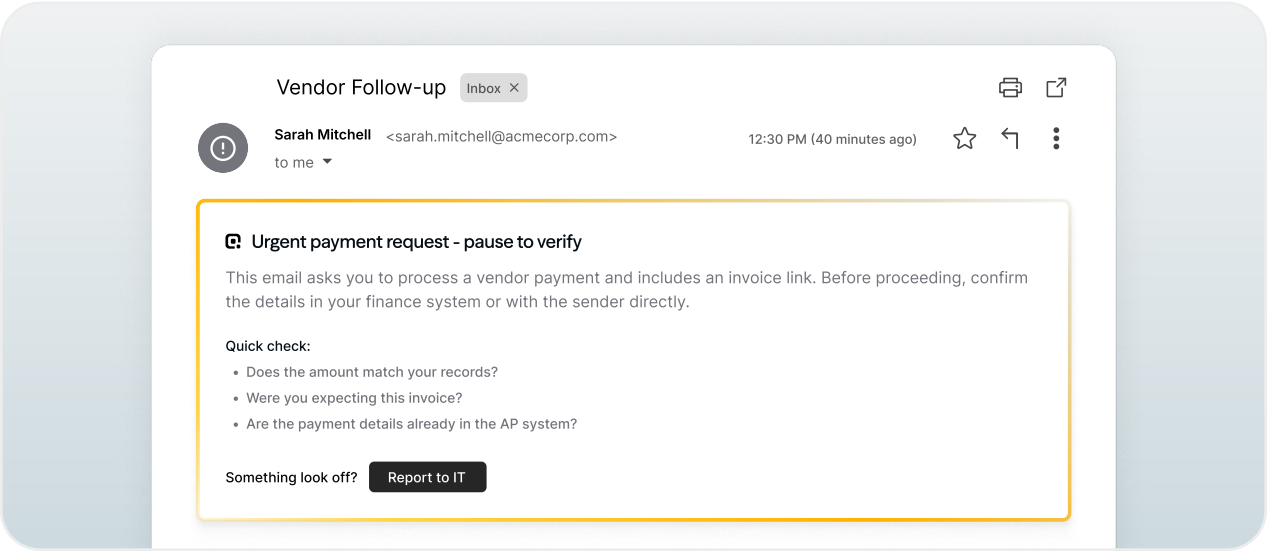
Pause any impacted transactions, access changes, or workflows pending verification. Limit the window for additional exposure without waiting for full confirmation.

3. Resolve

Complete a technical and factual investigation, close the attack vector, and remediate any downstream effects including access changes, financial transactions, or data exposures.

4. Strengthen

Feed incident findings directly into training updates, policy refinements, and simulation content. Every resolved incident improves the next response.



Severity Classification

Three tiers guide escalation and cross-functional engagement across all incident types in Section 3.

SEVERITY TIER	DESCRIPTION & TRIGGER
Tier 1: Potential	Suspicious content or behavior has been flagged. Deepfake involvement is unconfirmed. Security team investigates. No transactions are pending.
Tier 2: Confirmed	Deepfake involvement has been verified. No transaction or access compromise has occurred yet. Security, Legal, and the affected individual's office are engaged.
Tier 3: Active Impact	A deepfake led to a completed action: a wire transfer, credential disclosure, unauthorized access event, or data exposure. Full cross-functional response team activates.

The Control That Matters Most

Out-of-band verification is the highest-leverage control against deepfake attacks. The rule is simple: any request arriving by video, voice, or email gets confirmed through a separate channel before anyone acts on it. A known-good phone number. A direct callback. A confirmation from someone reached independently.

Rachel Tobac calls this practice “polite paranoia.” If a request comes in by email, call the sender directly to confirm it. If it arrives by phone, verify it through a second channel before acting. Teams that build this into standard operating procedure stop most deepfake attempts before they reach the transaction stage.

Two things matter equally when an incident is suspected: speed and accuracy. Moving quickly contains exposure. Moving on incomplete information creates a different kind of damage. The steps in each scenario below are sequenced to get both right.

SECTION 3

Five Incident Types: Detection to Resolution

The five incident types below cover the deepfake attack scenarios security leaders encounter most frequently. Each follows the four-phase structure. Recognition signals, response steps, and outcomes are specific to each attack type.

3.1 Executive Impersonation via Deepfake Video

A senior leader's video likeness and voice are cloned from publicly available content: earnings call recordings, conference keynotes, YouTube interviews, or LinkedIn video posts. The deepfake is delivered via a Zoom or Teams session, a pre-recorded message, or a combination of both. The request typically involves an urgent wire transfer, a policy exception, or a credential handover, framed to bypass standard approval channels.

The targeting of senior executives accelerated sharply through 2025. Deepfake fraud drained $1.1 billion from U.S. corporate accounts that year, triple the 2024 figure, and documented incidents had already quadrupled the prior year's total by mid-year.



Scammers cloned Italy's defense minister's voice and worked through the country's business elite, with at least one recipient sending nearly €1 million before realizing the call was synthetic.



A WPP CEO's likeness was replicated in a fabricated Microsoft Teams call, built from YouTube footage, requesting that a senior executive fund a new business venture.



Goldman Sachs executives were deepfaked in Instagram video ads promoting fictitious investment schemes.

The pattern across each case was consistent: multi-channel reinforcement, manufactured urgency, and a request that fit what the impersonated person might plausibly ask.

Recognition Signals

- The request arrives outside normal approval workflows or bypasses dual-authorization requirements
- Video or audio artifacts are present: unnatural blinking patterns, lip-sync inconsistencies, or audio compression irregularities
- The communication originates from an unfamiliar account or device even when the face and voice appear familiar
- The employee feels uncertain after the interaction and second-guesses the request before acting

Response Steps

Immediate (0-1 hour)

1. Pause any pending transactions or access changes requested during the communication.
2. Notify the security team and the impersonated executive's office through a known-good, out-of-band contact.
3. Preserve all available evidence: screen recordings, call logs, email threads, and device metadata.

Containment (1-24 hours)

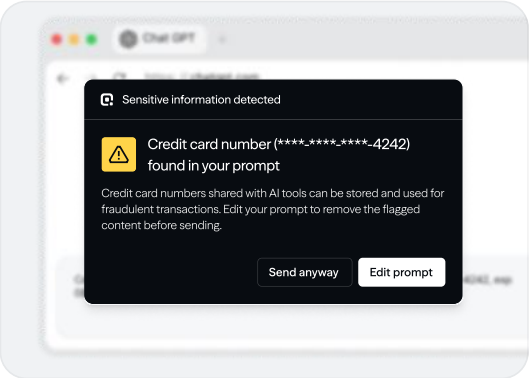
1. Conduct technical analysis on the video or audio content using deepfake detection tooling.
2. Review access logs for any downstream activity resulting from the request.
3. Engage Legal and Finance at Tier 2 or above. Issue a calm, factual internal advisory to employees in affected workflows.

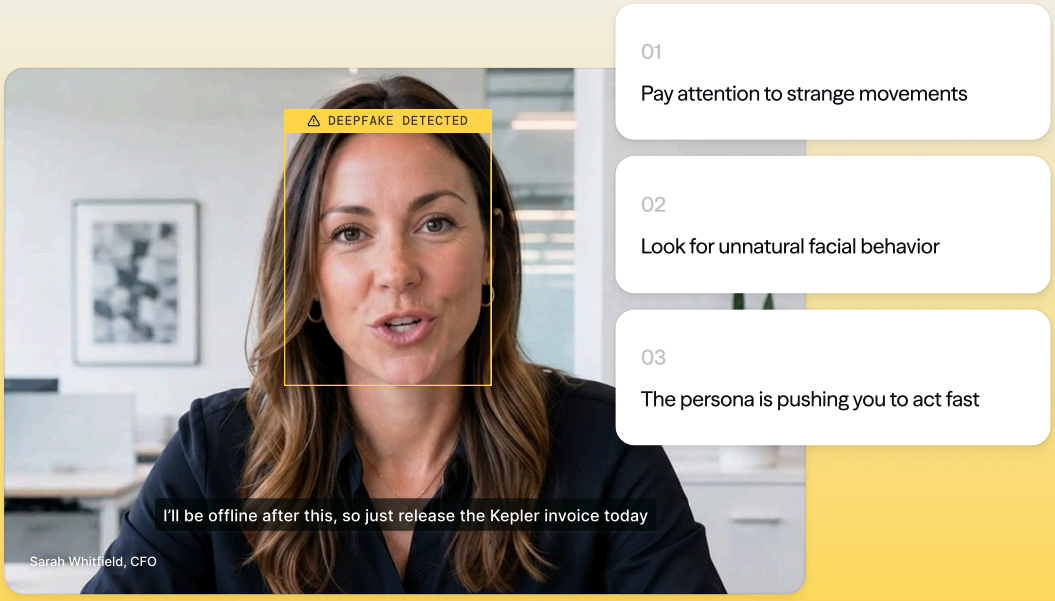
Resolution and Documentation (24–72 hours)

1. Confirm whether any financial or access-related damage occurred and initiate remediation as needed.
2. Document the attack vector, delivery method, and any indicators of compromise.
3. File regulatory or compliance notifications where required. Schedule targeted training for high-risk roles.

What good looks like

Teams where out-of-band verification is practiced regularly stop executive impersonation attempts before they reach the transaction stage. Adaptive's executive deepfake simulations build this habit across your organization before an attacker tests it.





3.2 AI Voice Phishing

Cloned voices are used in live calls or pre-recorded voicemails to impersonate executives, IT helpdesk staff, or vendors. The goals are consistent: extract credentials, capture MFA codes, approve financial transfers, or create a plausible pretext for a follow-up interaction. Siwei Lyu, a computer science professor at the University at Buffalo who directs its Media Forensic Lab, wrote in Fortune in December 2025 that voice cloning has crossed what he calls the "indistinguishable threshold." A few seconds of audio now produce a convincing clone, complete with natural intonation, rhythm, emphasis, and breathing. The perceptual tells that once gave away synthetic voices have largely disappeared. Earnings calls, podcast appearances, and conference talks provide that source material at scale for any publicly known executive. And attackers are increasingly targeting middle managers and controllers, where a three-second voicemail is enough.

Recognition Signals

- The caller requests credentials, MFA codes, or account access over the phone
- The audio carries subtle compression artifacts, a slightly mechanical cadence, or inconsistent emotional tone
- The caller is evasive when asked verification questions or applies pressure to skip standard protocols
- The callback number does not match known contact records for the person being impersonated
- The request arrives during a known absence, a board meeting, or another period when reaching the impersonated person through other channels would be inconvenient

Response Steps

Immediate (0–1 hour)

1. Instruct the employee not to act on the request until identity is independently verified.
2. Preserve any voicemail recording or call log. Verify the caller through a known-good contact number.

Containment (1–24 hours)

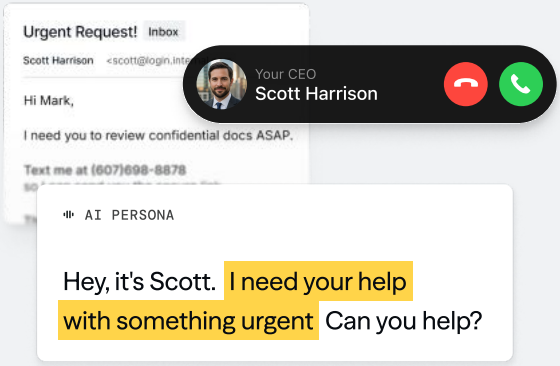
1. If credentials or MFA codes were disclosed, treat this as a full credential compromise immediately: revoke access, reset credentials, and review active sessions.
2. Analyze telephony logs and any access events linked to the impersonated individual.
3. Submit audio for voice cloning detection analysis where tooling is available. Notify stakeholders per the applicable severity tier.

Resolution and Documentation (24–72 hours)

1. Document the incident and update organizational threat intelligence with new indicators.
2. Brief employees in high-risk roles: Finance, Executive Assistants, and IT Helpdesk.
3. Review telephony verification policies for gaps and update where needed.

What good looks like

AI voice phishing is most effective against employees who have never heard a voice clone before. Adaptive's multi-channel simulation program includes voice phishing scenarios that let your team experience exactly how convincing a cloned voice sounds, before that experience happens in an uncontrolled setting.



3.3 AI-Powered Spear Phishing

Generative AI combined with open-source intelligence produces email campaigns that look indistinguishable from internal communications. The emails reference colleagues, active projects, current vendors, and recent company events. They arrive from addresses that pass casual inspection and use language that matches the organization's internal tone. Signature-based email filters were designed for a different era. Since the launch of ChatGPT, phishing attacks have grown 4,151 percent, and according to the FBI's 2024 Internet Crime Report, business email compromise alone cost U.S. organizations $2.9 billion last year. An estimated 40 percent of those emails were AI-generated.

Recognition Signals

- Emails reference internal project names, colleague names, or recent events in ways that would require insider knowledge or OSINT research
- Requests involve credential submission, file downloads, or wire transfers from unexpected senders
- Sender domain, email header metadata, or signature formatting carries slight inconsistencies on close inspection
- An above-normal volume of similar messages targets a specific team or individual within a short window
- An employee reports that something felt off, even if they cannot pinpoint why. That instinct should always be investigated

Response Steps

Immediate (0–1 hour)

1. Quarantine the suspicious email across the organization using email remediation tooling.
2. Determine scope: how many employees received the message and across which teams.
3. Identify any employees who clicked links, downloaded attachments, or submitted credentials.

Containment (1–24 hours)

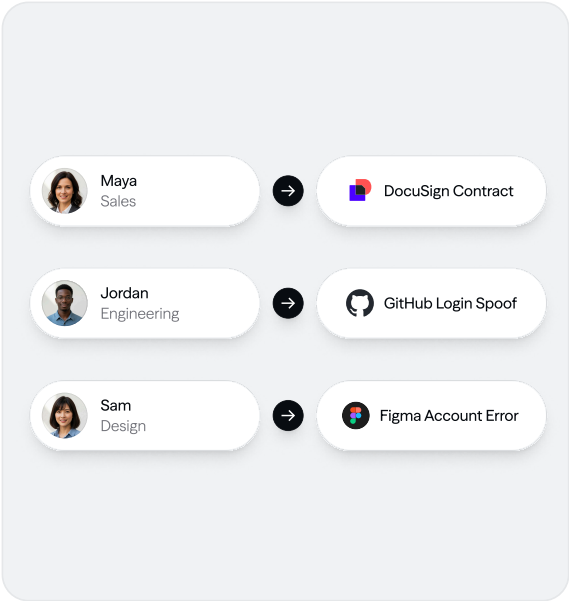
1. Reset credentials and revoke active sessions for any confirmed compromised accounts.
2. Analyze email headers, embedded links, and attachments for indicators of compromise.
3. Review access logs for lateral movement from compromised accounts.
4. Brief affected employees with clear, specific guidance on what happened and what steps they should take.

Resolution and Documentation (24–72 hours)

1. Update email filtering rules based on newly identified indicators. Tools like Adaptive's cloud email security platform can automate this step by continuously retraining on emerging AI-generated phishing patterns.
2. Feed campaign characteristics directly into future phishing simulation design.
3. Conduct a targeted training refresh for employees who interacted with the campaign.

What good looks like

Organizations running OSINT-driven phishing simulations with AI-personalized content see measurably lower click rates over successive rounds. After approximately twelve simulation rounds, detection success rates have moved from 34 percent to 74 percent in independent studies. Adaptive's simulation library is built to produce that kind of sustained improvement.



3.4 Deepfake-Assisted Hiring Fraud

AI-generated faces, cloned voices, and fabricated credentials are now used to pass hiring processes and gain employment at target organizations. Gartner projects that by 2028, one in four candidate profiles worldwide could be fake. Once inside, whether as employees, contractors, or vendors, the goal is access: to sensitive data, internal credentials, or a quiet foothold they can use later. Organized criminal groups have used this approach at scale against technology companies, with fabricated workers passing multiple interview rounds and clearing background checks. “Once they get the job, they attack the company,” as Brian Long has described. “They get access to internal systems, documents, payroll.”

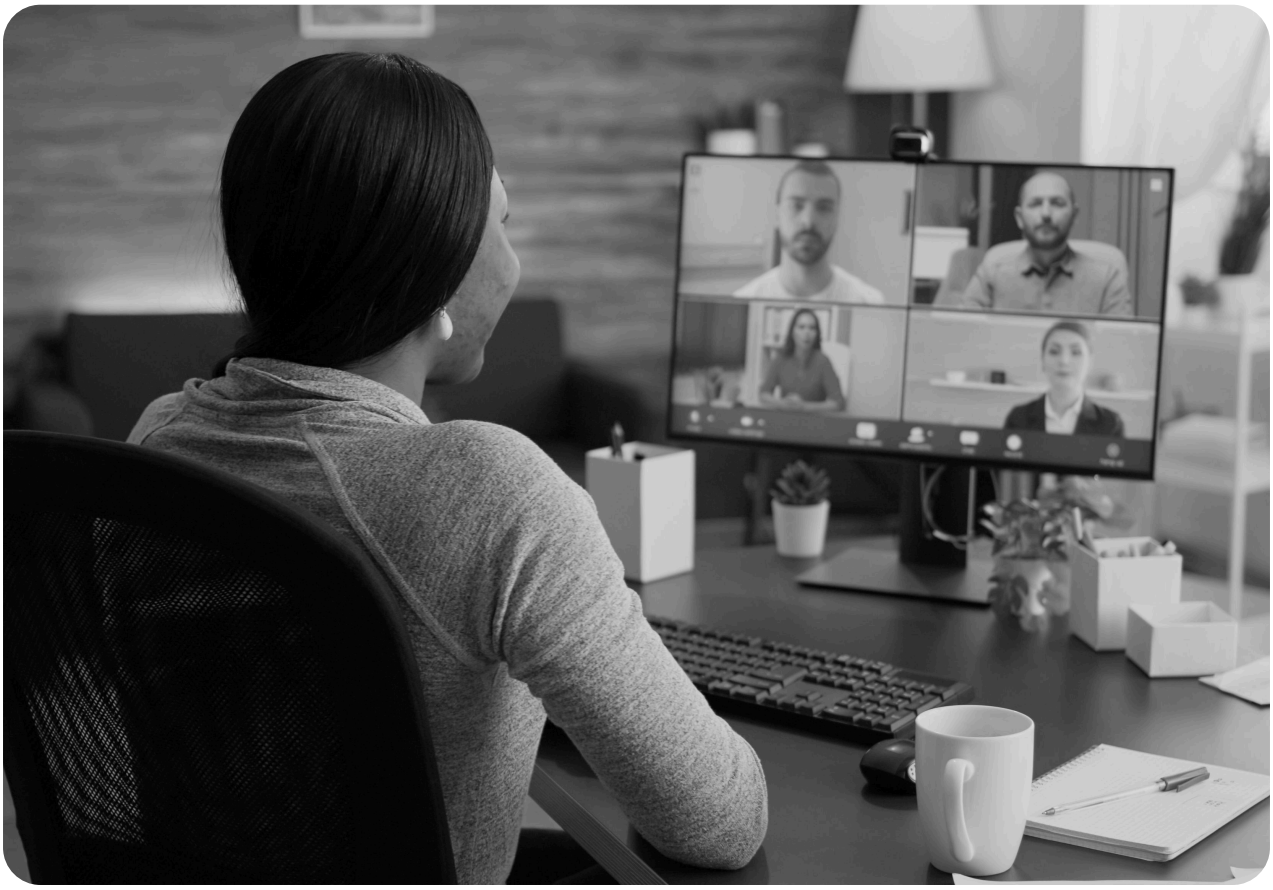
Recognition Signals: Pre-Hire

- Profile photos show artifacts associated with AI generation: unnatural symmetry, inconsistent lighting, or eye quality that does not match the rest of the image
- Technical interview answers follow a scripted pattern, with delayed responses and phrasing that sounds generated rather than recalled
- Video interview frames carry background inconsistencies or subtle face-to-voice mismatches
- References or credentials cannot be independently verified through the institutions cited

Recognition Signals: Post-Hire

- Access patterns fall outside the scope of the employee's stated role
- Requests for elevated permissions arrive earlier than expected in the onboarding timeline
- Data access or download volumes are atypical for the role level





Response Steps

Pre-hire detection (0–1 hours)

1. Pause or extend the hiring process for the flagged candidate pending a joint Security and HR review.
2. Conduct a reverse image search and independent identity verification through the cited credentials and institutions.
3. Document the candidate profile and retain all records for potential law enforcement reporting.

Post-hire containment (1–24 hours)

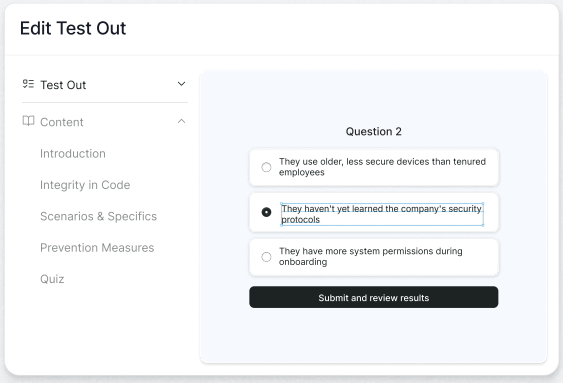
1. Revoke all access credentials and system permissions immediately.
2. Conduct a full audit of data accessed, downloaded, or transmitted by the individual.
3. Preserve onboarding records, device logs, and communication history as evidence.
4. Engage Legal on employment termination, regulatory reporting, and law enforcement notification.
5. Review any third-party relationships the individual introduced during their tenure.

Resolution and Documentation (24–72 hours)

1. Update hiring and onboarding verification procedures based on the specific gaps this incident exposed.
2. Brief HR and hiring managers on updated detection criteria with concrete examples.
3. Conduct a broader security review of systems the individual had access to.

What good looks like

Hiring fraud via deepfake is preventable with the right verification steps and trained interviewers. Adaptive's training library includes scenario-based modules for HR and hiring managers that build the pattern recognition needed to flag a fraudulent candidate before an offer is extended.



3.5 Reputational and Disinformation Attacks

Fabricated deepfake content featuring executives or employees is distributed internally or externally to damage trust, destabilize an organization ahead of a financial event, or manufacture a crisis. In July 2025, scammers used an AI-generated video of Ripple CEO Brad Garlinghouse to promote a fake 100 million XRP airdrop, timing the campaign to coincide with Ripple's high-profile legal win over the SEC. Ripple's own CTO had to publicly warn investors the video was synthetic. The content might show an executive making statements they never made, or attribute actions to an employee in ways designed to embarrass the organization. Distribution channels include social media, internal messaging platforms, email, and media contacts. The goal is disruption, and time is the attacker's advantage: the longer fabricated content circulates unchallenged, the more damage it does.

Recognition Signals

- Content attributed to a known executive or employee is unexpected, out of character, or contradicts publicly known positions
- Content surfaces through unofficial channels rather than corporate communications infrastructure
- Employees or external partners are forwarding or asking about a video or audio clip
- Media inquiries arrive about a statement the company has not made
- Internal messaging volume increases sharply around a piece of content

Response Steps

Immediate (0–2 hour)

1. Obtain and preserve a copy of the content as evidence before it is removed, edited, or shared further.
2. Convene a joint response team: Security, Communications, Legal, and the affected individual's office.
3. Conduct a preliminary technical assessment to determine whether the content is AI-generated, misattributed, or authentic material taken out of context.

Containment (2–24 hours)

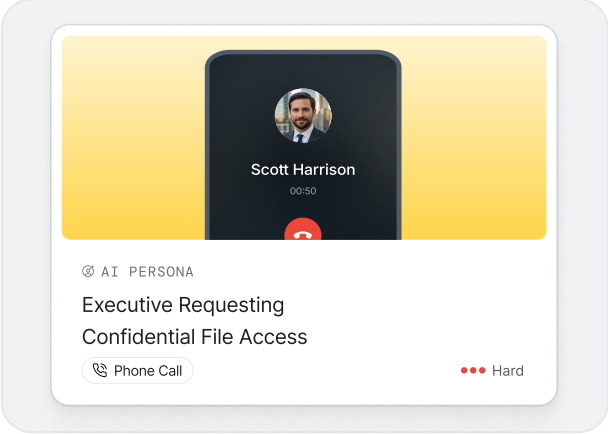
1. Draft internal communications to contextualize the situation for employees: factual, calm, and timely.
2. Prepare an external statement in coordination with Legal and Communications if the content is circulating publicly.
3. Submit platform takedown requests where applicable. Map the distribution chain to identify origin and propagation path.

Resolution and Documentation (24–72 hours)

1. Complete a formal forensic report on the content's technical characteristics and likely origin.
2. Notify regulators, law enforcement, or partners as appropriate.
3. Conduct employee-facing communication to close the loop. Review executive OSINT exposure monitoring to identify what source material may have been used.

What good looks like

Reputational attacks do the most damage when organizations have no prepared communications plan. Adaptive's executive exposure monitoring surfaces the OSINT data attackers use to build these scenarios, giving your security team visibility before an incident starts.



SECTION 4

Building a Proactive Defense Program

The response steps in Section 3 assume that an incident has already started. The organizations best positioned to handle deepfake attacks are the ones that have done specific preparation before that happens. Four areas determine how well a program performs under pressure.

Training That Reflects Current Attack Methods

Employees who have experienced a deepfake simulation in a controlled environment are better at identifying one under pressure. That finding holds consistently across independent research, and it aligns with what Adaptive observes across its customer base.

[The window for spotting deepfakes by eye has already closed] "I can't tell the difference these days. So, I think the time of looking for this type of deepfake by eye has long passed unfortunately."



DOMINIC FORREST
CTO OF IDENTITY VERIFICATION FIRM IPROOV

READ THE ARTICLE

Generic awareness modules that focus on spotting misspelled emails do not prepare employees for a live voice call impersonating their CFO or a video message requesting an urgent approval.

Role-specific training is the standard to target. Finance teams face different scenarios than HR teams, and both face different scenarios than IT staff. The training content and simulations should reflect those differences. The cadence should be regular enough that skills stay current as attack methods evolve.

Verification Protocols That Work Under Pressure

Out-of-band verification is the most consistently effective control against deepfake attacks. Any request involving a financial transaction, a credential change, or an access modification gets confirmed through a separate channel using a known-good contact method. The challenge is operationalizing that policy so it functions under the time pressure and authority pressure that deepfake attacks are designed to create.

> "They don't have to know anything about neuroscience or cybersecurity, but remember one code name and use it [when] necessary."



CHRISTOPHER HADNAGY
SOCIAL ENGINEERING RESEARCHER AND AUTHOR

READ THE ARTICLE

A challenge phrase used in high-stakes requests creates a verification step a deepfake cannot satisfy without organizational insider knowledge. Implementation effort is low. Protection value is meaningful. Escalation paths should be documented and practiced so employees know exactly who to call and through which channel when something feels wrong.

Technology Controls That Support Human Judgment

Deepfake detection tools integrated into video conferencing, email review workflows, and file analysis pipelines give security teams a technical layer alongside the human one. Neither is sufficient alone. Dr. Hany Farid, a professor at UC Berkeley and one of the leading researchers in synthetic media detection, has put it succinctly: detection technology will always lag behind generation technology. The tools improve, but so does the content they are trying to catch. Trained employees catch signals that automated tools miss. The combination is more reliable than either in isolation.

Executive exposure monitoring identifies the OSINT data available to potential attackers: public audio and video, organizational details, personal information indexed from professional and social platforms. Dynamic risk scoring identifies the employees most likely to be targeted based on role, tenure, and behavioral patterns. Knowing who is exposed and who is high-value to an attacker shapes where training resources and monitoring focus go.



Tabletop Exercises

Annual or semi-annual tabletop exercises that walk cross-functional teams through each incident type surface gaps that documentation reviews miss. Finance, HR, Legal, Communications, and Security each have a role in a deepfake incident response. A tabletop exercise confirms that those roles are understood and that the handoffs between teams function when it counts.

Readiness Assessment

Review the following questions with your team. A yes to all six means your program has the right foundation. A no in any of these areas identifies a specific gap worth closing before an incident surfaces it.

- Do employees know how to verify an unusual executive request through an out-of-band channel, and have they practiced it?
- Has your team run simulations that include AI-generated voice and video scenarios?
- Is there a documented escalation path for a suspected deepfake incident?
- Does your incident response plan include a cross-functional communications protocol?
- Are high-risk employees in Finance, HR, and Executive Assistant roles receiving role-specific training on a regular cadence?
- Is executive OSINT exposure being monitored proactively?

SECTION 5

Your Response Team and Post-Incident Review

Deepfake incidents rarely stay within the security team's scope. The response team that needs to function together spans six areas, and coordinating them in advance is the difference between a clean response and a reactive one.

Who Responds, and When

ROLE	PRIMARY RESPONSIBILITY	ENGAGES AT
Security Team / SOC	Detection, technical analysis, containment, evidence preservation	All Tiers
IT / Help Desk	Account lockdowns, credential resets, device forensics	Tier 2-3
Legal & Compliance	Regulatory notification, evidence handling, law enforcement liaison	Tier 2-3
Communications / PR	Internal messaging, external statements, media response	Tier 2-3
HR	Hiring fraud response, employee guidance, insider threat coordination	Tier 2-3
Finance	Transaction review, wire fraud response, vendor verification	Tier 2-3
CISO / Exec Sponsor	Cross-functional coordination, board reporting, major decisions	Tier 3



Pre-establishing communication templates for each audience, including internal employees, board members, regulators, and media, eliminates one of the highest-pressure decision points during an active incident. A communications plan that requires writing from scratch in the first two hours of a Tier 3 response creates avoidable mistakes.

Template 1: Internal Employee Notification

```
TO: ALL EMPLOYEES
FROM: [CISO NAME]
SUBJECT: IMPORTANT SECURITY NOTICE. PLEASE READ.
```

We are currently responding to a security incident involving synthetic media. You may receive communications that appear to come from senior leadership requesting urgent action on financial transactions, credential changes, or access approvals. Do not act on any such requests until you have verified them through [designated out-of-band contact method].

If you receive a suspicious request, report it immediately to [SOC contact / internal reporting channel]. Do not forward, screenshot, or share the content externally.

Further guidance will follow. Questions should go to [designated internal point of contact].

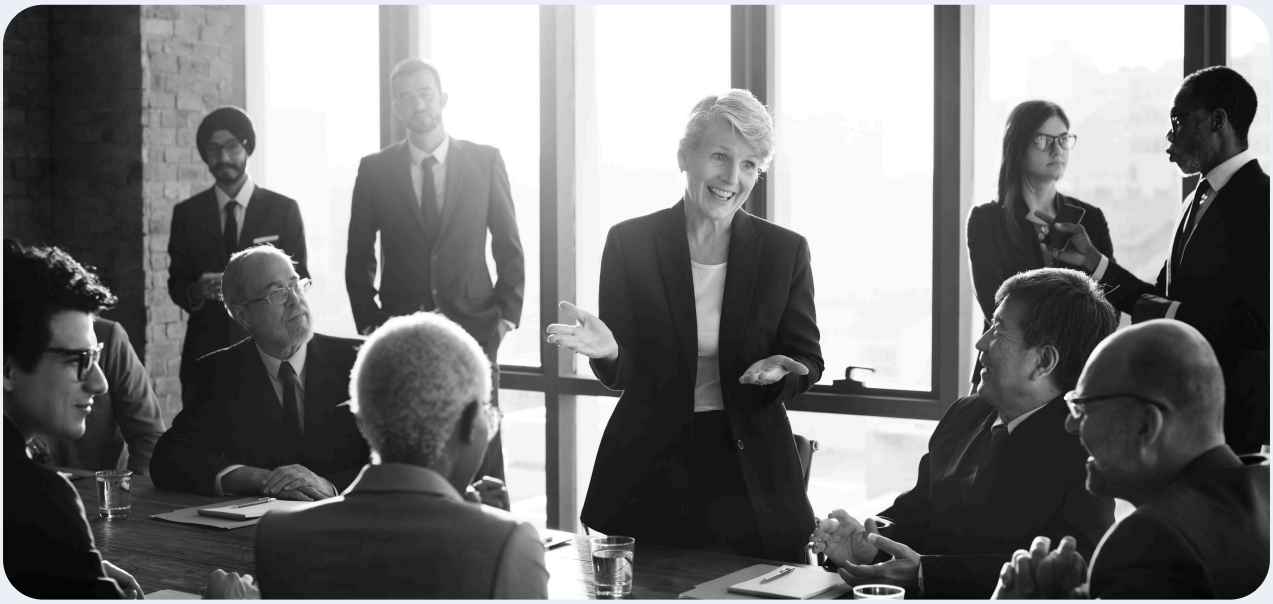
Template 2: Board Notification

TO: BOARD OF DIRECTORS
FROM: [CISO NAME]
RE: ACTIVE SECURITY INCIDENT. TIER 3 RESPONSE INITIATED.

On [date/time], our Security Operations team identified a [deepfake audio / video / synthetic email] incident targeting [brief description: e.g., a senior executive, a finance team member]. We have activated our Tier 3 incident response protocol.

Immediate containment steps are underway. Legal and Communications have been engaged. [Regulatory notification is being assessed / has been initiated] in accordance with our incident response obligations.
Current status: [Contained / Under active investigation / Monitoring for secondary exposure]

A full briefing will be provided at [scheduled time]. The primary point of contact for board inquiries is [name and contact].



Post-Incident Review

The most valuable thing a security team can do after a deepfake incident is turn it into a measurable improvement to their program. That requires a structured review process with specific timelines. Within 72 hours of containment, document what happened in chronological order, which response steps worked as planned, and where gaps appeared. The goal at this stage is factual capture while the sequence is fresh.

Within two weeks, conduct a structured debrief with each team that participated. Root cause analysis should cover how the attack entered the organization, what made it effective, and what specific changes to process, training, or technology would have changed the outcome. Produce a prioritized remediation list with owners and deadlines.

Dr. Josephine Wolff, a professor of cybersecurity policy at Tufts University's Fletcher School and author of You'll See This Message When It Is Too Late (MIT Press, 2018), has studied dozens of major breaches and documented a consistent finding: organizations that conduct structured post-incident analysis improve their security posture over time. Those that treat incidents as events to be closed rather than examined tend to encounter the same failure modes again. The breach itself carries limited value. What the organization does with it afterward is what matters.

Feed the findings back into the preparedness program. New attack patterns become simulation content. Updated indicators go into detection tooling. Lessons learned get documented in a section of this playbook that carries a revision date.

Metrics Worth Tracking Over Time

- **Mean time to detect (MTTD) for suspected deepfake incidents**
- **Mean time to contain (MTTC) after confirmation**
- **Employee reporting rate for suspicious content**
- **Simulation interaction rates by role over successive rounds**
- **Number of incidents reaching Tier 3 in a given year**

Improvement across those metrics over time is a concrete measure of program maturity, and a useful frame for board-level reporting on the human risk program.

SECTION 6

How Adaptive Security Partners With Your Team

The programs that perform best when a deepfake incident occurs are the ones that practiced simulations before it happened. That preparation spans four areas: training employees on what AI-generated attacks look and sound like, running simulations that match the scenarios your organization faces, surfacing the exposure data attackers use to build their approach, and giving your security team the operational visibility to act quickly when something gets through.



Simulations Built from Your Organization's Context

Adaptive's simulation library includes AI voice phishing, executive deepfake video, OSINT-driven spear phishing, and hiring fraud scenarios. Simulations are built using your organization's own context: executive names, reporting relationships, current projects, and internal language. That specificity is what produces the recognition instincts that hold up under pressure. Employees who have experienced a convincing deepfake simulation in a safe environment are measurably better at catching the same patterns when they appear outside of it.



Role–Based Training That Stays Current

Training is role-based and continuously updated. Finance teams, HR, Executive Assistants, and IT staff each receive content calibrated to the scenarios they are most likely to face. The content updates as attack methods evolve, so the training your team receives reflects what is happening now. There is no separate content library to manage and no manual update cycle to run.



Executive Exposure Monitoring

Adaptive surfaces the OSINT data already available to potential attackers: public audio and video, organizational details, personal information indexed from professional and social platforms. Knowing that exposure exists lets your team take proactive steps before it becomes the source material for an attack. For the executives most exposed, that visibility changes the risk calculation in meaningful ways.

Dynamic Risk Scoring and Phish Triage

Dynamic risk scoring quantifies individual and organizational risk based on behavior, role, and exposure. High-risk employees get targeted interventions. When a suspicious email reaches an employee, the phish alert button and AI-powered triage classify it, escalate it, and clean up organization-wide distribution within the same platform your team already uses for training and risk monitoring. No separate console. No additional integration work.

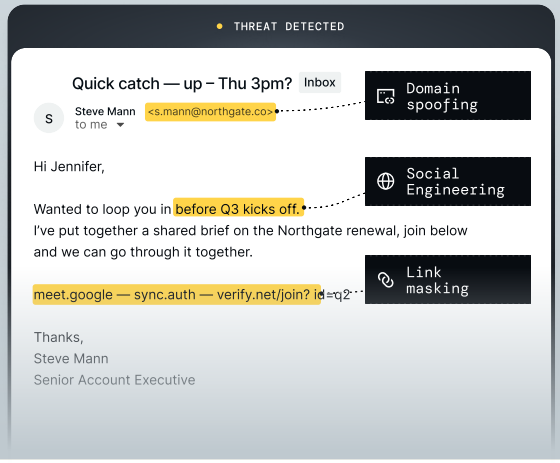
“The deepfake simulations were a huge wake-up call. They fundamentally changed how our employees think about social engineering.”



BEN PORTER
IT DIRECTOR, PODIUM

READ THE ARTICLE

The organizations that handle deepfake attacks well share a common characteristic: they treated human risk as a measurable, manageable program rather than a background condition. That is what Adaptive helps security leaders build.



Get Started

To see how Adaptive supports your deepfake response program, book a demo or take a self-guided product tour.

BOOK A DEMO

TAKE A TOUR