DARKTRACE

A CISO's Guide to Email Security

Defend your organization against phishing, business email compromise, and identity threats

Content

02	The 2026 email threat landscape
03	Key email-based threats to address
04	Looking to the past: The legacy approach
05	Case study: A real-world supply chain attack
06	What to look for in your email security vendor
07	A checklist for evaluating modern email security solutions
08	Why AI-powered email security is essential
08	Next steps

The 2026 email threat landscape

Today's organizations rely on a connected ecosystem of email, collaboration platforms, and SaaS tools to keep employees, customers, partners, and suppliers connected.

Business-critical activity and day-to-day collaboration both happen across multiple channels, creating a larger and more complex digital communications environment to secure.

At the same time, the nature of the threats targeting communication platforms is evolving. While attackers have always sought to exploit trust and normal business behavior, advances in technology have made it easier and more cost-effective to carry out sophisticated attacks at scale.

AI has accelerated this shift. Phishing messages are more polished, targeted, and easier to produce at scale – with AI-assisted phishing text doubling year-over-year.[1]

At the same time, attackers increasingly abuse trusted cloud services, compromise real user accounts, and launch these attacks beyond the inbox.

For CISOs, this means email security is no longer just a filtering problem. It's an identity, trust, data protection, and business-risk problem, one that requires defenses capable of adapting as quickly as attackers do.

73% say AI-powered cyber-threats are already having a significant impact on their organization.

- **State of AI Cybersecurity 2026**

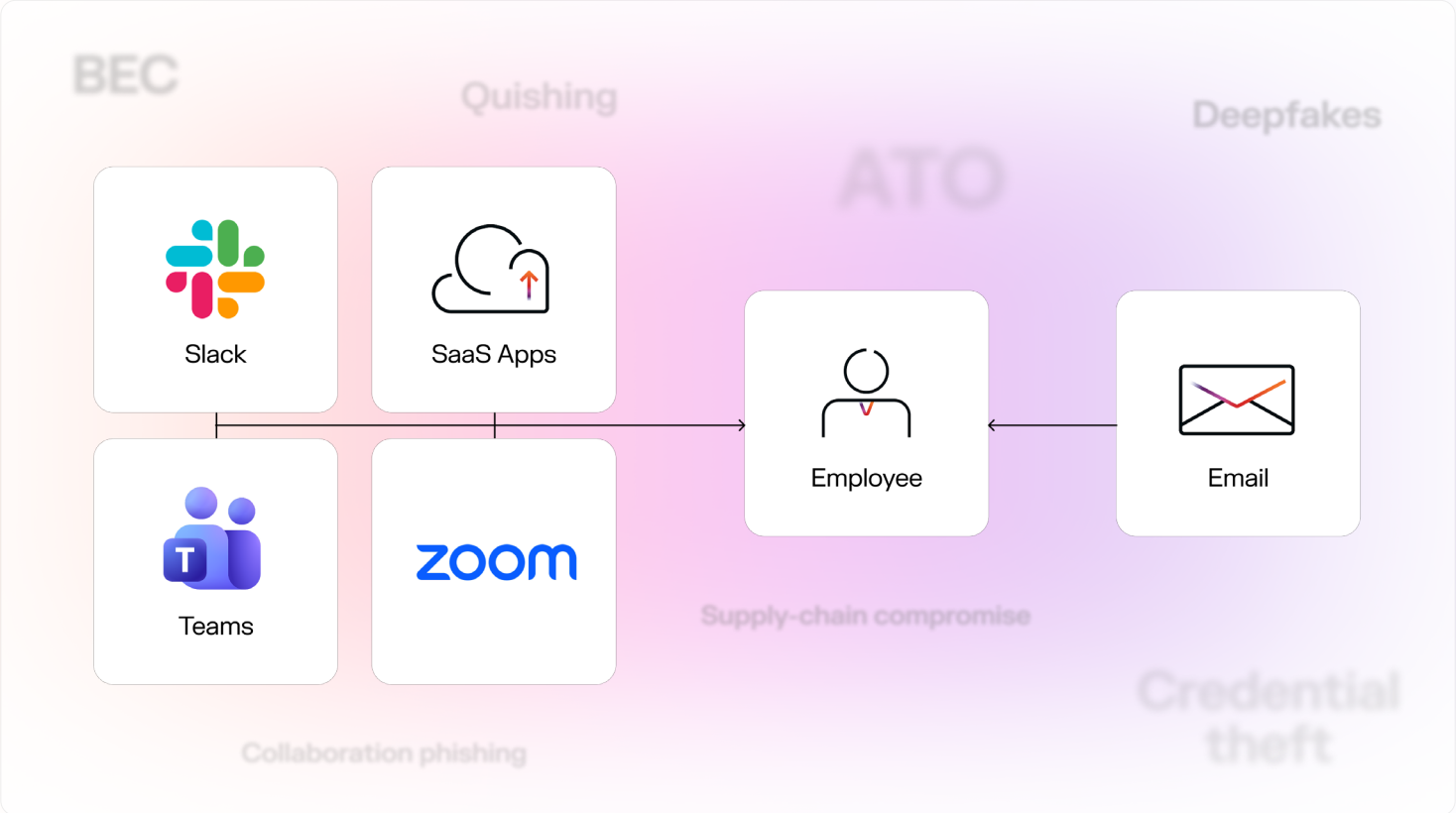

Figure 01: Threats now target the communication system, not just the inbox.

Key email-based threats to address

Credential theft

Social engineering attacks designed to harvest credentials, MFA tokens, or sensitive data, often serving as the first step toward account takeover, lateral movement, or ransomware

Business Email Compromise (BEC)

Highly targeted impersonation attacks that exploit trusted relationships to trigger fraudulent payments, data exfiltration, or unauthorized changes, frequently without payloads.

AI-driven social engineering

Generative AI enables attackers to produce more personalized, and grammatically adept messages at scale, reducing traditional detection signals and employee awareness.

Account takeover (ATO)

Compromised email accounts leveraged as trusted launch points for internal fraud, outbound phishing, supply chain attacks, or deeper access into cloud environments.

Supply chain & third-party impersonation

Attackers exploiting vendor, partner, and customer relationships to conduct cross-organization fraud and large-scale campaigns.

Abuse of legitimate infrastructure

Malicious content delivered via trusted cloud services (such as file-sharing platforms and collaboration tools) to evade reputation-based defenses and security controls.

Collaboration tool phishing

Social engineering conducted through platforms like Microsoft Teams, Zoom, or Slack, where users are less cautious and security visibility may be limited.

QR code (“quishing”) attacks

QR codes embedded in emails redirecting victims to malicious sites on mobile or unmanaged devices, bypassing traditional link analysis and endpoint protections.

Deepfake-enhanced impersonation

The use of AI-generated audio, video, or imagery to reinforce executive impersonation and high-stakes fraud scenarios.

AI-enhanced attacks

- **ranked #1**
  emerging enterprise risk, 3 quarters running[2]

~30%

- **of all data breaches**
  now involve a third party or supply chain compromise.[3]

When it comes to something like phishing emails, training on how to spot these is important but we simply cannot put the focus on humans to spot these well-researched, targeted email attacks.

- **CEO**
  Transport & Logistics

Looking to the past: The legacy approach

Key takeaways:

01 Trained on other organizations' data

02 Struggles with novel, targeted social engineering

03 Requires constant tuning and manual effort

04 Treats human risk as a separate problem

For years, email security has been built around the same core idea: identify what's known to be bad and block it. Traditional solutions (including SEGs) are trained on external threat intelligence that is designed to catch threats that have already happened elsewhere.

In practice, this means they need a Patient Zero. An attack must be seen, analyzed, and turned into a rule before it can be stopped. That delay leaves organizations exposed, especially to novel and socially-engineered attacks like Business Email Compromise (BEC) and Vendor Email Compromise (VEC), where there may be no obvious malicious indicators.

Attack-centric tools also tend to operate at the inbox level, often missing subtle signs that an account has been compromised or that a broader campaign is unfolding across SaaS and collaboration channels. Because these systems typically evaluate these channels separately, security teams are left to manually connect the dots, typically through SIEM or SOAR platforms that weren't designed for real-time, cross-channel coordination.

Other vendors have adopted a behavioral approach with AI that goes beyond rules and signatures. However, not all behavioral approaches are built the same. Some rely on aggregated datasets, cross-customer learnings, or generalized models of "normal" behavior to identify suspicious activity. While this can improve detection, it does not provide a complete understanding of what is expected, unusual, or risky for a specific organization.

This gap is also reflected in how organizations approach human risk. Without a clear understanding of normal user behavior, security teams rely heavily on inadequate awareness training as a primary line of defense. Intel from these programs is rarely fed back into strengthening email defenses.

The result is a noisy, high-maintenance security stack. False positives disrupt business workflows, analysts spend hours maintaining policies and reviewing alerts, human reporting lacks accuracy, and defenders are constantly reacting to yesterday's attacks. As attackers increasingly micro-target users, abuse trusted infrastructure, and move laterally across communication channels, it's clear that legacy, attack-centric email security is no longer fit for purpose.

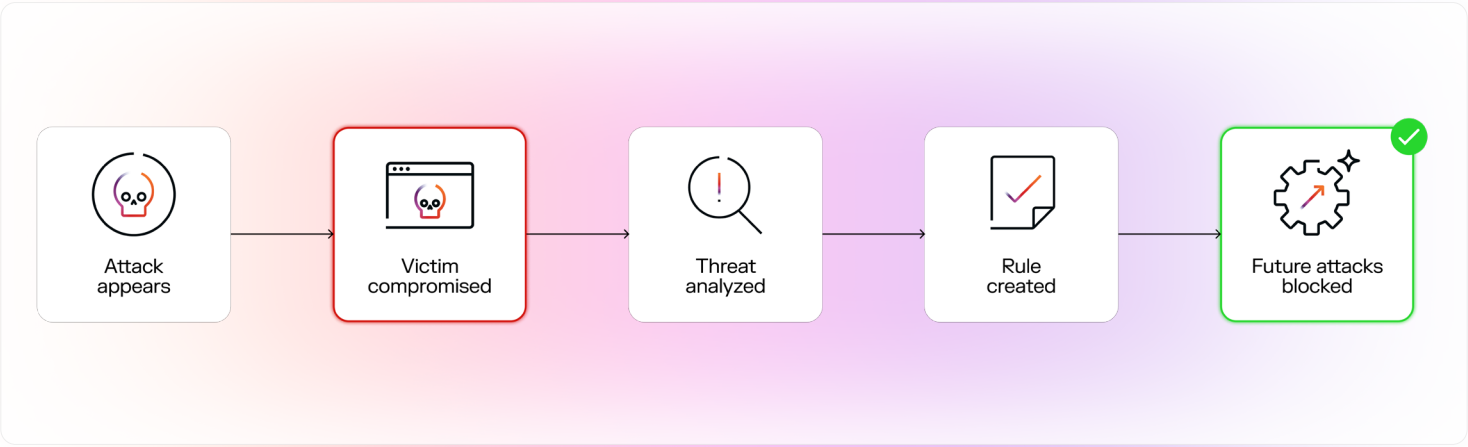

Figure 02: The Patient Zero problem

■ Case study

A real-world supply chain email attack

Let's look at an example of an attack analyzed by the Darktrace global SOC involving trusted vendor compromise. This attack would have bypassed an attack-centric email security solution.

Company profile: Large enterprise with extensive third-party vendor relationships and established email communication patterns

Attack overview:

Darktrace identified multiple suspicious vendor emails over a two-week period in 2025. Although the messages appeared to come from trusted partners and passed traditional security checks, several anomalies signaled malicious intent:

- Emails carried file-sharing links hosted on legitimate services that were previously never used in communications with this customer.
- Several senders were unknown to the network or behaved out of character for known contacts (e.g., sending to multiple undisclosed recipients).
- Message content and link behavior deviated from normal patterns, despite benign subject lines and trustworthy domains.

70% of malicious emails now pass DMARC.
Darktrace Annual Threat Report 2026

What traditional security missed

These emails were assigned a Spam Confidence Level of 1, meaning Microsoft's native filters did not mark them as spam or malicious, and they passed SPF/DKIM/DMARC authentication.

The embedded links directed users to credible file-sharing sites that then redirected to fake Microsoft login pages designed to harvest credentials – a multi-stage phishing tactic.

Why it failed:

- Trusted sender assumption: The SEG treated messages from known vendor domains as inherently legitimate, missing that the vendor accounts were compromised.
- Reputation-based checks can't see intent: Because links were hosted on reputable platforms and authentication checks passed, the SEG did not flag them.
- No behavioral context: Traditional defenses couldn't detect out-of-character communication or deviations from normal email patterns.

This attack demonstrates how attackers can exploit trust and legitimate infrastructure to slip past legacy defenses, and how behavior-based anomaly detection can expose such threats even when conventional indicators fail.

Read the blog for more detail.

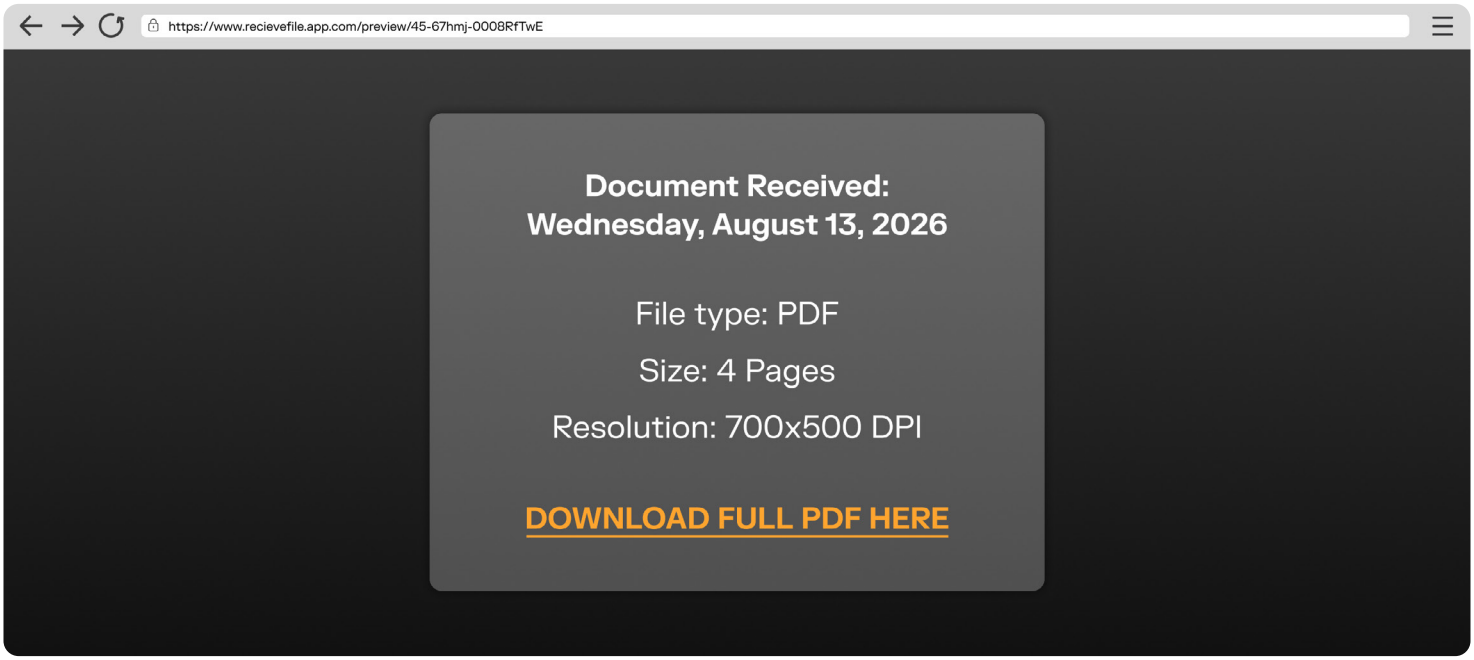

Figure 03: A legitimate file sharing site used in phishing emails to host a secondary malicious link.

What to look for in your email security vendor

Email security is no longer about blocking malicious messages, it's about understanding behavior, intent, and risk in real time, specific to your organization.

As attackers use AI to enhance phishing success and abuse identities, organizations should be rethinking what they need from their email security stack.

01 AI trained on your business

Instead of relying on other organization's threat data, leading solutions use AI to understand normal patterns of communication for each organization, team, and user, and identify when something deviates in a meaningful way. Enabling you to catch threats unique to your organization, supply chain, and collaboration patterns – threats that solutions trained on someone else's data will never see.

02 Email can no longer be treated as a standalone problem

Today's attacks involve account takeover, data loss, and collaboration tools like Microsoft Teams or SaaS applications. Effective email security vendors take a 360° view, correlating signals across inboxes, identities, endpoints, and cloud platforms to detect threats earlier and respond more precisely – without relying on third-party acquisitions or integrations that create siloes and miss context.

03 For AI security vendors, transparency is key

Most solutions force a choice between automation, but without the ability to see or influence decisions, or transparency, but the security team has to write and maintain the rules. Organizations deserve autonomous protection that doesn't require a full-time operator, with depth and adjustability for teams that want to engage when they choose to.

04 Make employees part of the solution, not the problem

Humans don't just need protecting, they can form an active part of your defense. With the right guidance and training, employee behavior can help strengthen your security posture, not weaken it.

05 Finally, the relationship between security teams and email security tools needs to change.

In 2026, CISOs should expect solutions that reduce manual effort, act autonomously when appropriate, and integrate cleanly with native platform protections.

A checklist for evaluating modern email security solutions

Use this checklist to evaluate whether an email security vendor is built for today's threat landscape or anchored to the past.

- [ ] Can the solution detect novel and highly targeted attacks (including BEC, VEC, and AI-driven social engineering) without relying on prior examples?
- [ ] Is the AI trained on your organization's data? Or does it rely on aggregated datasets and threat intelligence? Where does your data go? Is it centralized with other people's data?
- [ ] Is detection based on behavior and context, or primarily on rules, signatures, and pattern-matching?
- [ ] Can you get transparency into every AI decision? Or is it "black box" AI?
- [ ] Does it provide real-time visibility and response across communication channels, including collaboration tools (e.g. Microsoft Teams), and SaaS applications?
- [ ] Does it monitor outbound communications to prevent accidental or malicious data loss and insider-driven risk?
- [ ] Does the AI learn continuously how your organization communicates, or does it require extensive upfront tuning and ongoing rule maintenance?
- [ ] Does it integrate with native email and cloud security controls to enhance existing protections rather than duplicating cost and complexity?
- [ ] Will it meaningfully reduce analyst workload, improving signal quality, automating response, and streamlining investigations?
- [ ] Does it improve user-reported phishing accuracy, using employee training and feedback as a signal to strengthen detection over time?

Why AI-powered email security is essential

Modern email threats don't follow rules. Effective email security in 2026 requires multi-layered AI – not a single model of an LLM added on top of legacy controls.

Effective solutions use multiple forms of AI to understand how an organization normally communicates and to identify meaningful deviations in real time. By learning behavior rather than relying on other organizations' attack data, these systems can detect novel and highly targeted threats without waiting for prior examples.

To be effective, AI analysis must extend beyond the inbox. Today's attacks involve compromised accounts, lateral movement, and collaboration tools. By correlating signals across email, identities, and cloud applications, the AI can detect and respond to threats wherever communication occurs.

When applied correctly, AI reduces complexity rather than adding to it. Autonomous decision-making and workflow automation lower false positives, improve the quality of user-reported phishing, and free security teams to focus on higher-impact risks.

Next steps

The shift towards AI-powered attacks abusing trusted relationships, compromised accounts, and legitimate infrastructure exposes the limits of legacy email security.

Tools built to recognize known threats struggle when attacks seem trustworthy on the surface and unfold across identities and collaboration platforms. Protecting the organization now depends on analyzing behavior, context, and intent with deep AI understanding of your unique organization, not just filtering out malicious messages.

For security leaders and boards, the priority is clear: adopt email security that adapts continuously, strengthens native protections, and reduces operational friction – allowing organizations to preserve trust, productivity, and resilience in their email operations.

Key takeaways for security leaders

- Modern email threats exploit **identity and trust**, not just technical weaknesses
- Legacy and broad behavioral tools lack visibility into **your specific risks**
- **AI trained on your org's data** is required to detect novel attacks
- Email security must cover **inbound, outbound, lateral, and collaboration channels**
- Effective solutions r**educe workload while improving protection**
- **Integrated employee training** can strengthen your security posture

About Darktrace / EMAIL

Email and collaboration security built for modern threats

Darktrace / EMAIL is designed to stop the attacks that legacy email security misses, those that abuse trust, mimic normal business behavior, and move across inboxes, identities, and collaboration channels.

Unlike approaches that rely on models built across many organizations, Darktrace's AI learns exclusively from your organization's activity. It builds an understanding of how your people communicate and collaborate, enabling it to detect subtle deviations without relying on rules, signatures, or prior examples.

- **58% of threats detected by Darktrace pass through other email filters**
  Identifies advanced BEC, supplier compromise, and payload-less social engineering that blend into normal business workflows.
- **Blocks novel threats up to 13 days faster**
  AI understands your unique email and messaging patterns to stop attacks before traditional tools can recognize them.
- **Automates response across email, identity, and SaaS**
  Reduces manual triage, alert fatigue, and console switching through coordinated, autonomous action.
- **Protects your brand, messaging, and collaboration tools end to end**
  Provides visibility and protection across users, domains, and collaboration tools (including Zoom, Slack, and Microsoft Teams) not just the inbox.
- **Improves signal quality by 60%**
  Empowers analysts and employees with AI-driven triage and contextual guidance, leading to better end-user reporting and reducing SOC investigations.
- **Unifies email defense and security awareness training**
  Delivers measurable behavior change through adaptive, bi-directional learning for employees.
- **Fastest deployment in the industry**
  Responds to threats up to **30x faster**, delivering value in days, not months.
- **Proven and trusted**
  Recognized as a Leader in the 2025 Gartner Magic Quadrant for Email Security and a Gartner Peer Insights Customers' Choice 2026, with a 4.8★ rating from customers.

Want to learn more?

View the product →

Calculate your potential ROI with Darktrace / EMAIL →

About Darktrace

Darktrace is a global leader in AI for cybersecurity that keeps organizations ahead of the changing threat landscape every day. Founded in 2013 in Cambridge, UK, Darktrace provides the essential cybersecurity platform to protect organizations from unknown threats using AI that learns from each business in real-time. Darktrace's platform and services are supported by 2,300 employees who protect nearly 10,000 customers globally. To learn more, visit darktrace.com.

North America: +1 (415) 229 9100 Europe: +44 (0) 1223 394 100 Asia-Pacific: +65 6804 5010 Latin America: +55 11 4949 7696

darktrace.com | info@darktrace.com