REPORT

proofpoint

The 2025 Data Security Landscape

Introduction

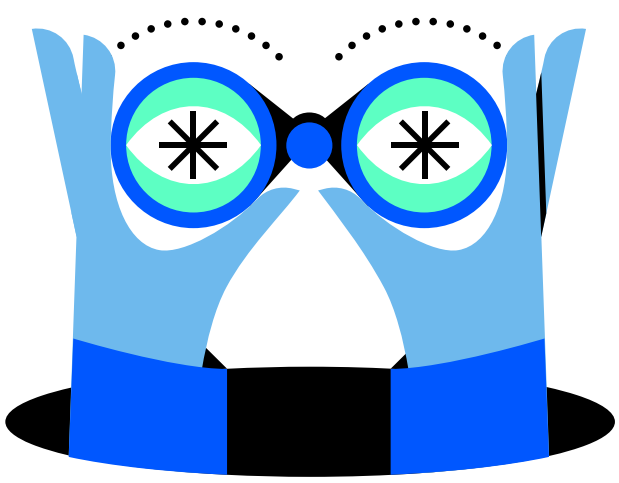
Data security is at a major inflection point. For some time, it has required an increasingly layered approach: one that accounts for human behavior, data sprawl and the adoption of generative AI (GenAI). But the emergence of AI agents—happening at an unprecedented rate—is giving rise to a new, more complex agentic workspace where humans and agents work together. With this shift, new risks are emerging. The story of data security is entering a new chapter.

Information sprawl and unprecedented growth in volumes of enterprise data have become major challenges. Organizations store data across complex hybrid and multicloud environments. And as amounts of information continue to grow, it's becoming harder to identify, classify and protect it. To make things worse, today's knowledge workers increasingly create "shadow" data: information that's created and stored outside the visibility of IT teams. As a result, insider risks have become magnified.

Adding to the complexity, AI has altered the data security landscape like few technological advances before it. Public GenAI tools such as ChatGPT, enterprise AI tools such as Microsoft Copilot and custom large language models (LLMs) have increased the potential for sensitive enterprise data to become exposed. What's more, organizations across all industries are hurrying to deploy AI agents to transform their business workflows. Companies must prepare today for the increased data security risks posed by the new agentic workspace.

The Proofpoint *2025 Data Security Landscape* report explores the current state of enterprise data security: the frequency and leading causes of data loss incidents, as well as today's biggest data security risks and challenges. It also looks to the future by examining how organizations are already evolving—or planning to evolve—their data security programs to reduce insider risk, streamline security operations and facilitate safe AI adoption. These data security and insider risk goals have taken on added urgency as the agentic workspace becomes a reality.

Key findings



29%

of organizations say their **volume of data has increased by 30%** or more in the past 12 months.

48%

say **limited visibility into risky user behavior** is a top challenge for securing sensitive data.

85%

of organizations experienced a **data loss incident** in the last 12 months.

50%

say a leading benefit of a unified data security solution is **enabling safe and productive use of AI in their organizations.**

1%

of users were responsible for 76% of data loss events.

65%

are already **using security capabilities enhanced by AI** to classify data.

Enterprise data loss is still a people problem

Reflecting the breadth of data security risks and challenges organizations now face, our 2025 survey reveals that data loss incidents are still widespread. And despite the increasing presence of AI in modern organizations, careless insiders are still the leading cause.

Data loss incidents are pervasive

Data loss incidents are affecting all countries and industries. Of the 1,000 organizations polled in our survey, 85% experienced a data loss incident in the last 12 months. This was even higher in the U.S, U.K. and UAE, where over 90% experienced at least one incident.

What's more, most organizations suffer these incidents regularly. Respondents experienced a mean number of 11 data loss incidents in 12 months—nearly one a month. More than half (52%) had between one and 10 incidents in this period. Another 18% saw between 11 and 20. An unfortunate—or perhaps unprepared—5% suffered 31 or more.

The human element remains the weakest link

Clearly, enterprise data loss incidents are still far too frequent. But what are the causes? Our survey finds that these incidents are still primarily a people problem. From a range of common causes—spanning human behaviors and technical vulnerabilities—we asked respondents to choose all that led to data loss incidents in the past 12 months. Careless employees or third-party contractors are cited by 58%—the leading cause. In addition, 42% point to compromised users, while almost a third (32%) attribute malicious insiders.

58%

say that careless employees or third-party contractors were a cause of data loss incidents.

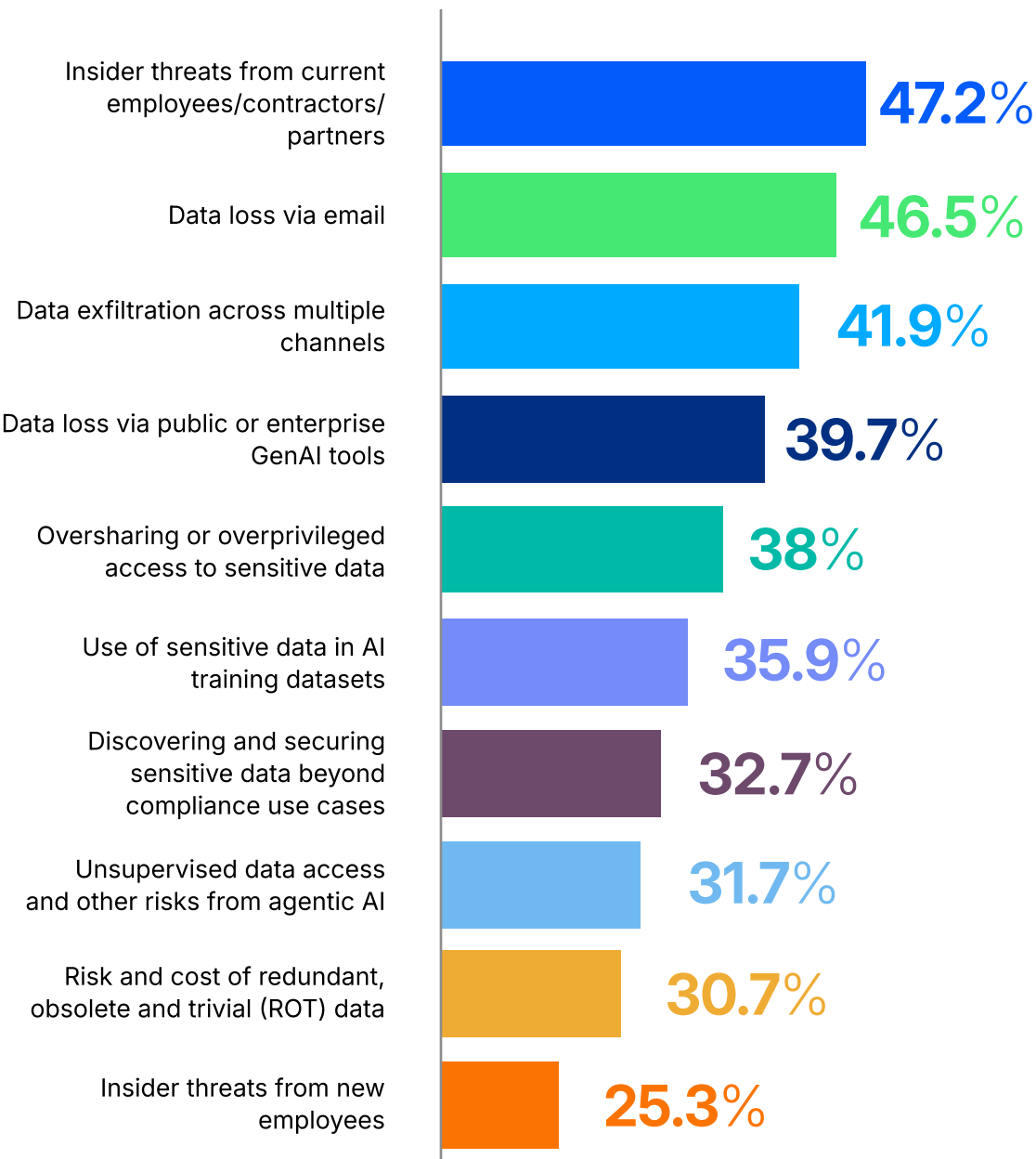
Proofpoint platform data also reveals a striking insight: on average, just **1%** of users are responsible for **76%** of data loss events. The makeup of this 1% can shift over time as employees leave an organization, others join and business dynamics change. However, this statistic underscores the need for behavior-aware security solutions that respond to real-time user actions and context. Those might include capabilities such as dynamic user risk scoring and adaptive insider risk policies. Clearly, using human-centric solutions to target the riskiest users can have profound security impacts.

Data security today: complex and evolving

To assess the current state of enterprise data security globally, we surveyed organizations on topics such as the amounts of data they manage, their top data security risks and challenges, the numbers of data security tools they use and their efficiency in resolving data loss incidents. Their answers were illuminating.

Ten different types of risk are cited by 25% or more of our respondents as top concerns. The findings highlight major contributors such as risky user behavior, data growth and loss of sensitive data. While insider threats are the top concern globally, there are interesting variations by country. In the U.S., Australia and Italy, data loss via email is the top risk. For other countries, AI risks are the biggest concern. In Germany, for example, 50% point to data loss via public or enterprise GenAI tools. In UAE, 46% cite the use of sensitive data in AI training datasets.

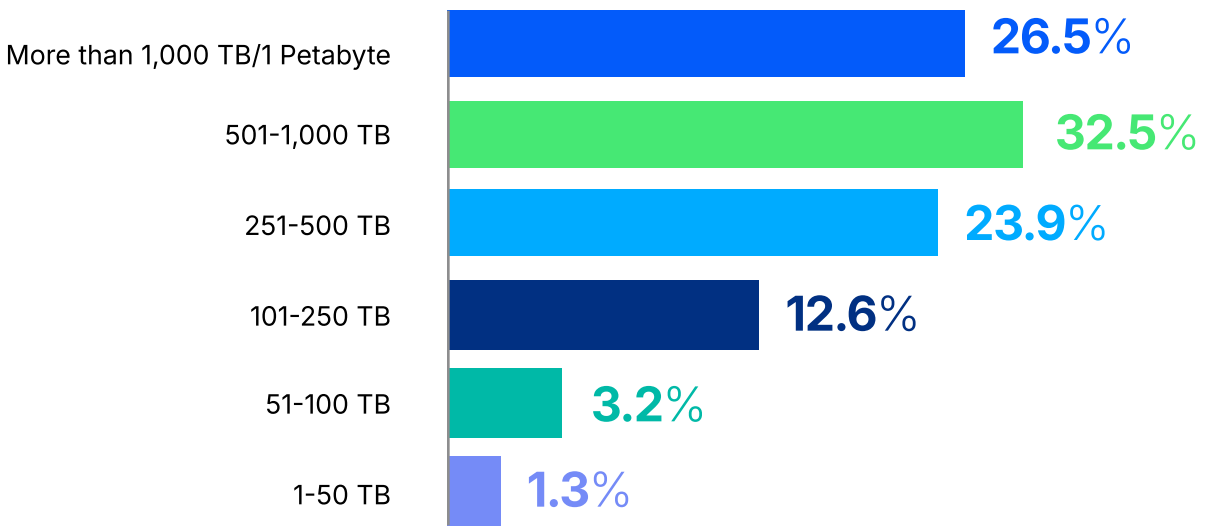
Which of the following are your top data security risks? (Select up to five.)



No end in sight for enterprise data growth

Most organizations are managing a *lot* of data. In fact, 59% of companies with 1,000 or more employees handle over 500 terabytes (TB). More than a quarter (27%) manage more than a petabyte (PB). For companies over 10,000 employees, the percentage managing over a petabyte increases to 41%. These are staggering amounts that would have been hard to imagine even a few years ago—and understaffed security teams are struggling to keep up. Drivers of this growth include data-intensive business processes, lower costs of data storage (especially on cloud platforms) and the demands of AI models and agents. These factors will continue to drive data volumes upward.

How many terabytes (TB) of data does your organization have worldwide?

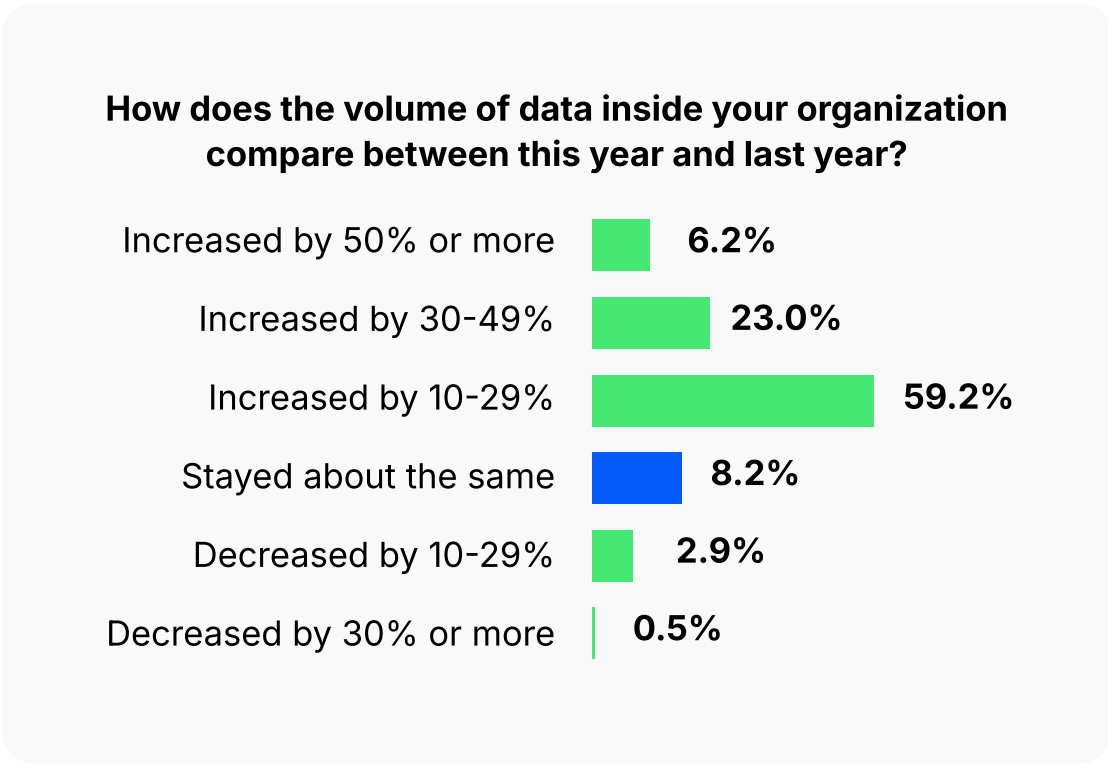


Countries around the world are managing data at different levels, based on local business requirements, digital transformation and AI adoption. For example, 70% of organizations in Brazil and 68% in UAE are managing over 500 terabytes (TB)—far above the global average. Interestingly, in Japan, less than 50% of organizations have 500 TB or more. This might be due to factors such as stricter controls around data collection, prevalence of legacy systems and a robust disposal process for unused data.

Quantities of data are also increasing at an unprecedented rate, adding to the pressure on security teams. For example, 59% saw the amount of data they handle grow between 10% and 29% over the past year. Another 23% saw even faster growth of between 30% and 49%.

41%

of companies with over 10,000 employees manage more than a petabyte (PB) of enterprise data.



Companies recognize the negative impacts of this data growth. For 46%, data sprawl across cloud and SaaS applications is a top data security challenge. And 31% say redundant, obsolete and trivial (ROT) data causes significant risk.

46%

say that data sprawl across cloud and SaaS applications is a significant data security challenge.

Proofpoint platform data validates these concerns. It shows that 27% of storage volumes on Amazon Web Services, Microsoft Azure and Google Cloud Platform are abandoned. This means that for every petabyte of data an organization stores in the public cloud, 270 TB might be unused. These unneeded data stores drive up cloud storage and backup costs. They also expand the attack surface without delivering value, as well as increasing compliance and legal risks.

More sensitive data, more risk

As amounts of enterprise data skyrocket, quantities of sensitive data are on a similar trajectory. In fact, 55% say that as much as 21–40% of their data is sensitive. In the U.K., the percentage of organizations rises to 66%.

55%

say that 21-40% of their data is sensitive.

Organizations are struggling to keep these huge amounts of sensitive data secure. For example, 38% cite oversharing or overprivileged access to sensitive data as a top risk. And Proofpoint platform data shows that 16% of sensitive files in Microsoft 365 are shared widely. This means they are shared company-wide or have public links. For an average organization, this can amount to tens of terabytes of overshared sensitive data.

This oversharing creates significant exposure—and serious compliance risks. For example, if an organization has 500 TB of data and 25% is sensitive, that's 125 TB of information requiring protection. With many organizations managing even more than that, the scale of the challenge is immense.

Underlining the problem, Proofpoint intelligence about the most frequent risks in public cloud environments has a recurring theme: storage containers in Amazon Web Services, Microsoft Azure and Google Cloud Platform that contain sensitive data, but which are publicly exposed due to overprivileged access or misconfiguration.

AI in the enterprise: powerful but perilous

The rapid adoption of AI—particularly GenAI and AI agents—is rapidly transforming the modern workplace, promising huge gains in productivity and operational efficiency. But according to our respondents, these benefits also bring new data security risks.

For example, 44% say that lack of visibility and data security controls for GenAI tools is a significant challenge for securing sensitive data. In addition, 36% are concerned about the use of sensitive data in AI training.

44%

say that lack of visibility and data security controls for GenAI tools is a top challenge for securing sensitive data.

What's more, AI agents, which often require broad access to enterprise data, can act as privileged superusers. Without proper design and governance, they pose serious risks of data leakage and misuse. Reflecting this, nearly a third (32%) are concerned about unsupervised data access and other risks posed by agents.

Unsurprisingly, countries vary in their concerns about AI adoption and are at different maturity levels. In both Germany and Brazil, 50% say that data loss via public or enterprise GenAI tools is a top risk—10% above the global average. In UAE, 46% say that use of sensitive data in AI training is a leading risk, also roughly 10% above average. As organizations roll out AI to drive innovation and productivity, AI governance will be critical for establishing guardrails and educating employees.

Blind to the inside: a top risk, but with limited visibility

What are the different types of insider risk?

An insider can be an employee, contractor or partner: anyone in a position of trust. All insiders pose risks. But an insider becomes a threat when they misuse their access to negatively impact a company's critical information or systems.

There are three types of insider risk: →

Careless insiders have good intentions but make mistakes. These insiders often take the fastest path to accomplish a task, which might mean unknowingly bypassing company policy. For example, an engineer at Samsung accidentally copied sensitive internal source code into ChatGPT in 2023.

Malicious users are motivated by personal gain and intend to harm the organization. Examples include data theft, espionage, sabotage and fraud. In 2025, Rippling, a workforce management platform, sued its competitor, Deel, over insider data theft and espionage.

Compromised insiders have their credentials stolen by threat actors looking for access to an organization's data and systems. In 2022, the LastPass breach was caused by an external threat actor who stole a developer's credentials and proceeded to steal source code.

As organizations adopt AI, they must be vigilant about risks from insiders—whether careless, malicious or compromised. For example, 47% globally see insiders as a top data security risk—the highest among the risks we surveyed. In Japan and the U.K., this rises to 59%. Given that data theft is one of the most frequent insider use cases, it's not surprising that 42% are concerned about data exfiltration across multiple channels, including email, endpoints and the web. And rightfully so. Malicious insiders are determined: if they aren't successful stealing data via one channel, they will often try another.

Compromised insiders—those whose credentials are stolen by threat actors—pose significant risks. Proofpoint platform data shows that 64% of tenants had at least one compromised account in the last year. Of those compromised accounts, nearly half (47%) saw suspicious post-access data activity. Not surprisingly, our research indicates that threat actors target valuable data, both structured and unstructured. The file types with the most suspicious post-compromise activity were .xlsx (17%), .docx (14%) and .pptx (11%).

Adding to this, most data security teams are not confident about their ability to handle rogue insiders. Almost half (48%) say that limited visibility into risky user behavior is a top challenge for securing sensitive data.

48%

say that limited visibility into risky user behavior is a top challenge for securing sensitive data.

Email: a hotspot of insider-led data loss

Email continues to be a major vector for insider-driven data loss. Globally, 47% of organizations cite email data loss from careless and malicious insiders as one of their biggest risks, second only to insider threats more generally. In the U.S., this percentage rises to 59%. Examples of email data loss include emails accidentally sent to wrong recipients (misdirected emails), emails with incorrect attachments and intentional exfiltration of sensitive data.

Email data exfiltration is when someone inside an organization emails sensitive data to a personal or unauthorized account. This account might be in a freemail, education or personal domain. Common motives include taking valuable or confidential information to a competitor, stealing it to sell to bad actors, or a desire to sabotage an organization.

Proofpoint conducts email data loss assessments to help companies identify email exfiltration. Some real examples of exfiltrated data types are shown.

Financial

- Annual holdings reports
- Client lists
- Documents labelled as confidential
- Investment schedules
- Password spreadsheets
- Portfolio reviews
- Profit sharing calculations
- Signed contracts

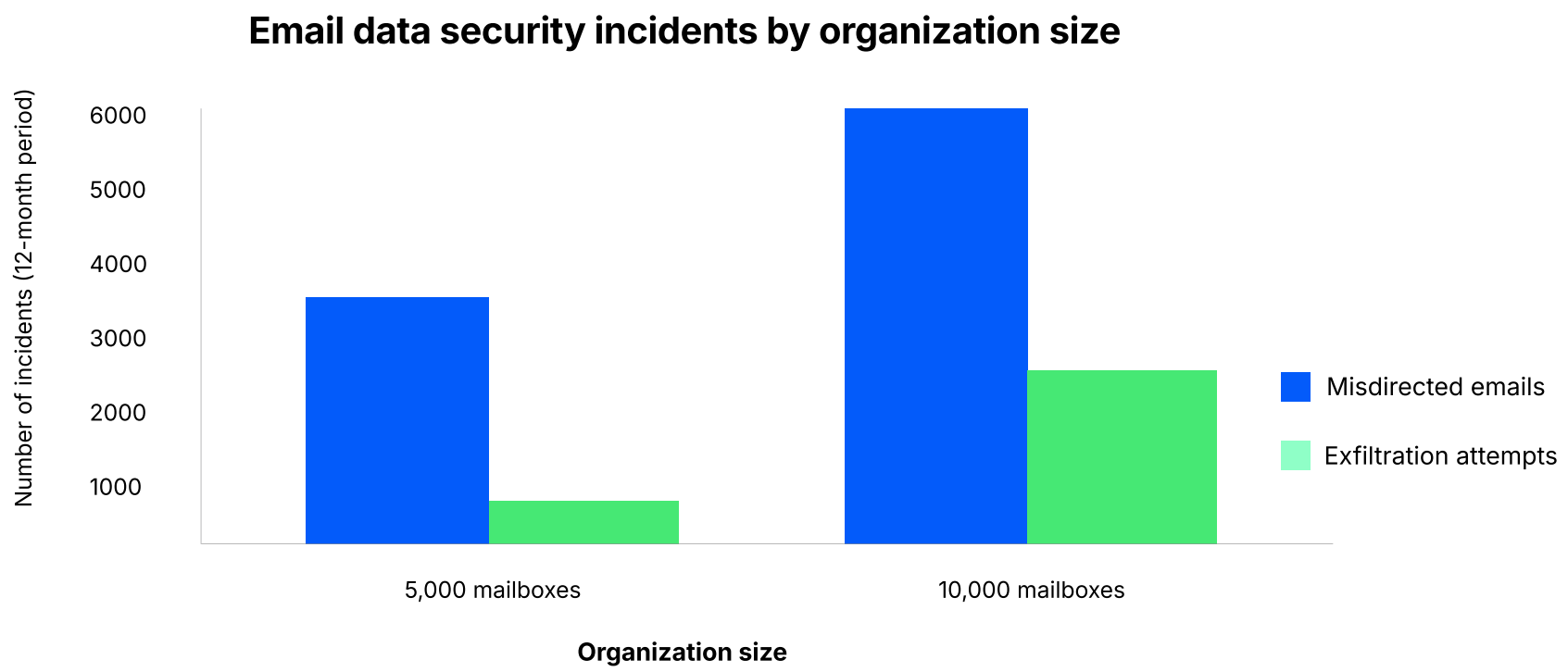
Healthcare

- Analgesic charts
- Anesthesia reports
- Medical reports
- Monthly practice reports
- Patient data
- Pharmacy passwords
- Provider agreements

Legal

- Client briefings
- Client waiver applications
- Draft proposals
- Meeting notes
- Password lists
- Political candidate notes
- Settlement documents
- Work authorizations

Again, Proofpoint platform data validates these concerns. It shows that in a 12-month period, an organization with 5,000 mailboxes sees an average of 3,607 misdirected emails and 632 attempts to exfiltrate sensitive data. For organizations with 10,000 mailboxes, these numbers scale dramatically to 6,567 misdirected emails and 2,566 exfiltration attempts. More than 60% of data exfiltration attempts include documents that contain sensitive data. Of further concern is that 39% of all misdirected emails also contain attachments.



Security teams: understaffed and overburdened

Data volumes and sprawl are growing and insider risks are being magnified by AI adoption. In the face of these risks, enterprise data security teams are coming under increasing strain. Indeed, over a third of respondents (35%) say that lack of qualified security personnel is a significant challenge. And 31% rely on just part-time resources to support their DLP tools.[1] A successful approach to data security goes far beyond technology to include people and process. With limited dedicated expertise, a data security program will struggle to protect the business. To help, organizations might consider engaging external service providers to deliver the skill sets needed to optimize their data security investments.

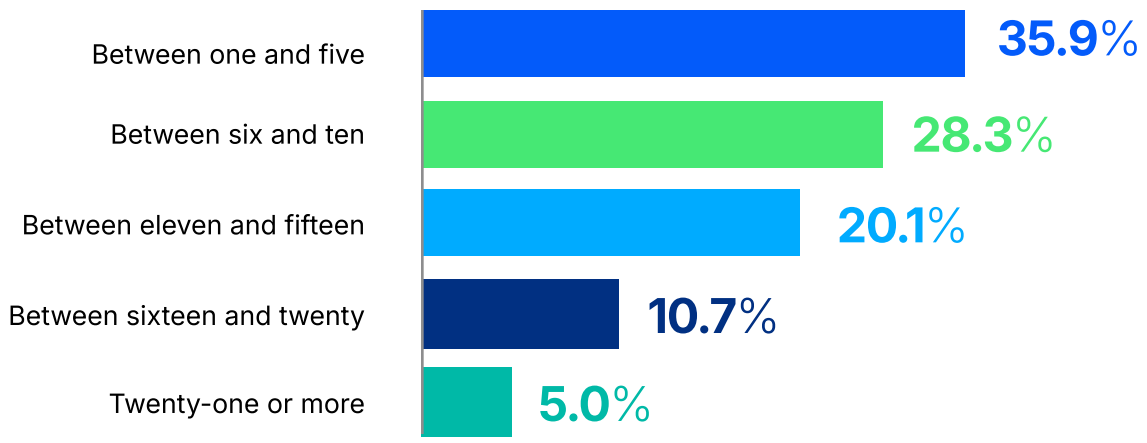
Given their resource constraints, security teams can ill-afford the time it takes to investigate and resolve frequent data loss incidents. Over a fifth (21%) say their teams take four to five days to resolve typical incidents. More concerning, a further 21% require **one to four weeks**. As well as being a prolonged drain on security team resources, these lengthy investigations keep sensitive data exposed and extend opportunities for threat actors to cause serious harm.

In light of these challenges, organizations recognize the burden of not having a unified data security solution to streamline their security operations. The fifth most cited data security challenge (41%) is having too many data security tools to administer and integrate.

1. Proofpoint. *Voice of the CISO report*. 2025.

Underlining this, the majority of respondents (64%) say they work with six or more data security vendors to secure sensitive data. Worse, 36% juggle more than 10.

How many vendors do you work with today to secure your sensitive data?



The complexity of administering so many tools increases operational overhead and reduces efficiency for already overstretched teams. This tool sprawl also requires analysts to pivot between multiple tools, compromising visibility and prolonging incident responses.

21%

of security teams take one to four weeks to resolve data loss incidents.

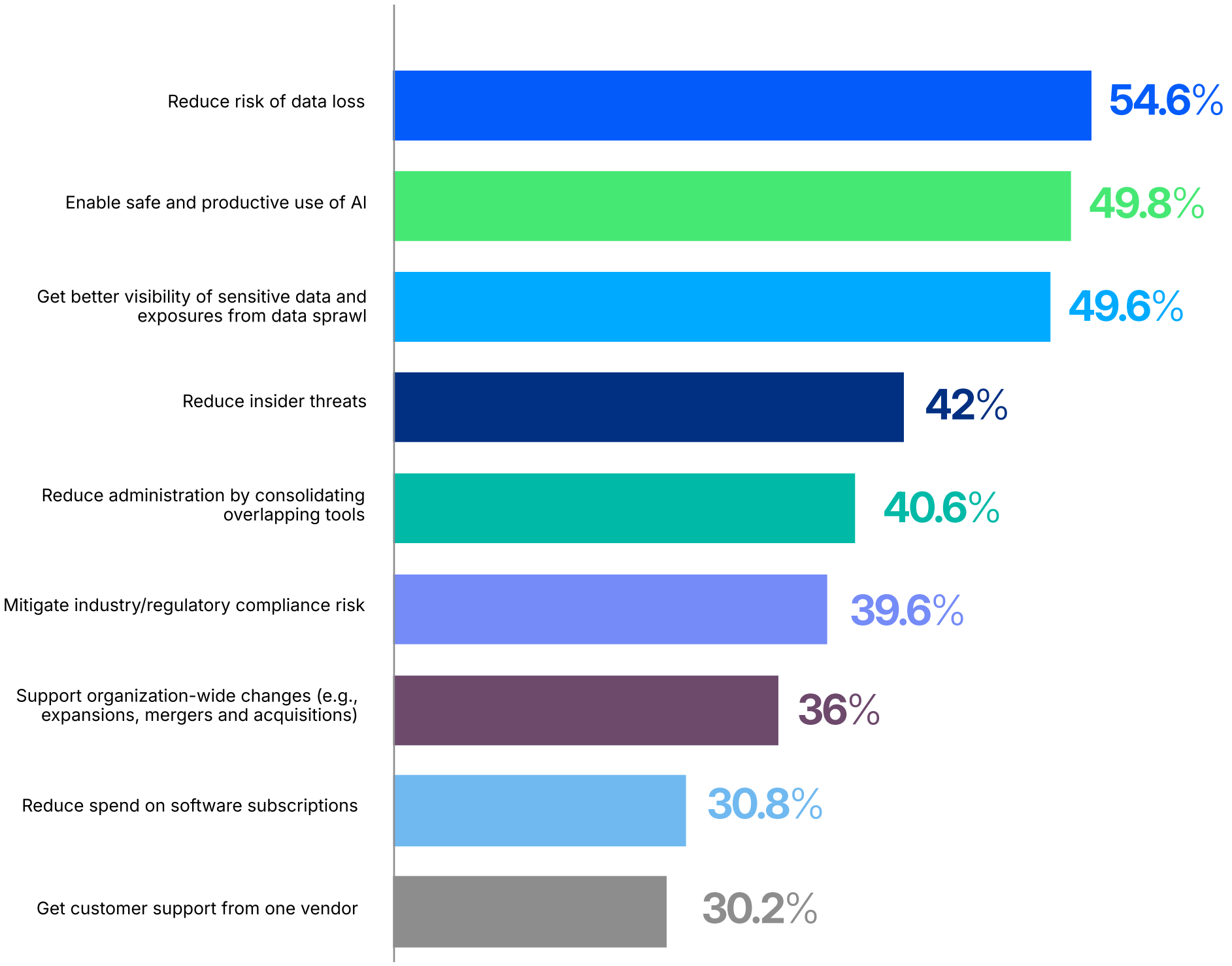
Looking ahead: holistic, AI-driven data security

Data security is at a pivotal juncture. Given the demands of fast-evolving workspaces comprised of humans and agents, how are security leaders thinking about the future? How will they streamline and strengthen their data security and insider risk programs while still enabling their organizations to embrace AI-driven productivity and innovation?

Unified data security is a must

Given the complex sets of risks and challenges they now face, security leaders increasingly recognize the need for unified data security solutions. These solutions must provide the visibility and controls to protect their organization's crown jewels, proactively detect insider risk and minimize data exposure. And to do so efficiently and quickly.

Which of the following are, or would be, the most significant benefits of a unified platform for data security? (Select up to five.)



Not surprisingly, 55% expect a unified solution to reduce data loss risk—the most cited benefit. In addition, 50% see a leading benefit as better visibility of sensitive data and exposures created by data sprawl. Clearly, security leaders recognize the role AI and automation can play in helping them understand where their data is, who has access to it and how it's configured.

With successful AI adoption an undoubted priority for organizations, unified data security is seen as a key enabler. Half of respondents (50%) say one of the biggest advantages of a unified solution is enabling safe and productive use of AI. In the U.S., which was an early adopter of AI, this increases to 64%.

50%

see one of the biggest benefits of a unified solution as enabling safe and productive use of AI in their organizations.

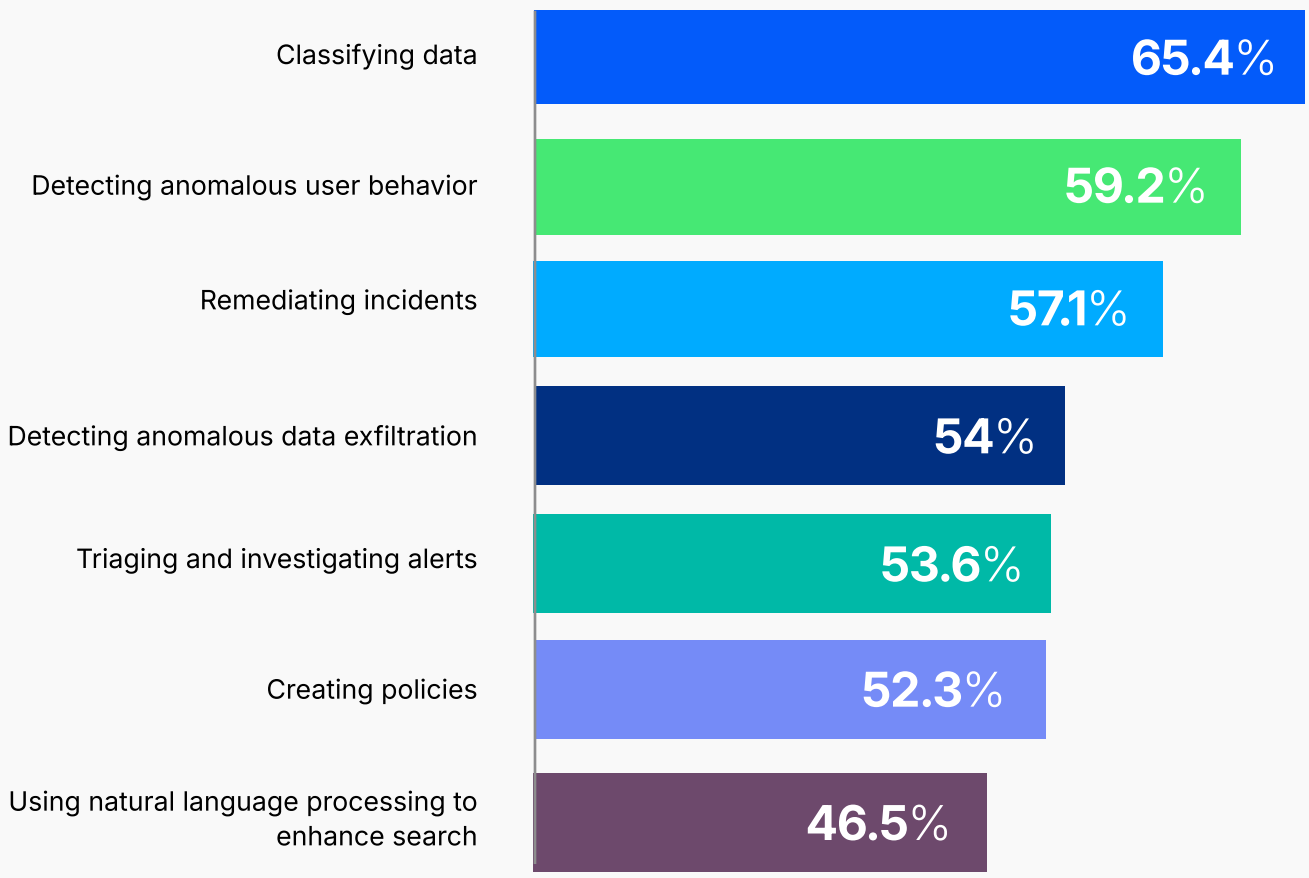
With human error cited as the leading cause of data loss, organizations also see how a unified solution can streamline security operations, improve visibility and reduce these events. For example, 41% say a big advantage of a unified solution would be consolidating overlapping tools to reduce administration overhead. And 42% recognize the benefit of reducing insider threats.

AI: the force multiplier

It's clear from our survey that organizations are worried about risks from AI. But there's a fascinating flip side: most also see the opportunity to harness AI as a force multiplier for their data security programs. In fact, for key data security priorities, over 46% have already deployed AI enhanced capabilities.

Those already using AI are using it in various ways to gain efficiencies and accelerate detection and response. Nearly two-thirds (65%) are using AI to classify data, while 59% are using it to detect anomalous user behavior. As security teams embrace AI to meet—and get ahead of—the challenges of the agentic workspace, they also see it as a way to enhance their security postures and address current pain points, such as reducing false positives.

Which of the following data security capabilities enhanced with AI (including agentic AI) are currently in use?



Conclusion

The data security landscape is defined by complexity, scale and accelerating technological change. Insider threats, data growth and sprawl, GenAI adoption and the rapid embrace of AI agents are converging to create unprecedented risks. The emerging agentic workspace—where both humans and agents interact with sensitive data—poses complex new risks that traditional data security and insider risk technologies simply weren't designed to manage.

Fragmented toolsets, limited visibility and overburdened security teams only compound these problems. The result is that organizations are experiencing frequent data loss incidents, sometimes needing weeks to resolve them. With those prolonged exposures come serious and potentially costly legal and compliance risks.

Yet, the path forward is clear: unified data security solutions that empower security teams, protect sensitive data and mitigate insider risk. These solutions must use AI-driven capabilities and enable behavior-centric security strategies. Organizations must move beyond patchwork solutions and embrace solutions that reduce security risk, lower operational cost and enable business agility.

To keep pace with emerging risks from both humans and agents, companies should have processes and controls in place to:

Use AI to identify and classify data unique to the business to prevent unwanted exposure and exfiltration

Visualize risk and prioritize incident response by tracing data lineage across channels

Implement adaptive controls that remediate access quickly for high-risk users

Ensure secure design and governance of AI systems to minimize risks of data leakage and misuse

Prioritize and continuously evaluate AI-enhanced capabilities, including agents, that help detect, triage and remediate data security incidents

Identify skill gaps that might stall optimization of your data security investments and consider options for leveraging specialized expertise

To learn how Proofpoint's unified, AI-powered data security solution can be the central pillar of your organization's future-focused data security and insider risk program, visit proofpoint.com/defend-data.

Methodology

To compile the *2025 Data Security Landscape* report, we:

- Surveyed 1,000 organizations across 10 countries and numerous industries.
- Drew on data collected from deployments of the Proofpoint human-centric security platform worldwide. Our platform data is one of the largest, most diverse data sets in the cybersecurity industry.

Survey data

Proofpoint partnered with cybersecurity market research firm, CyberEdge Group, to develop the 20-question survey instrument, to localize the survey instrument into non-English languages, to host the survey, to facilitate survey completions by qualified research participants and to analyze survey results. All respondents are IT security professionals employed by a commercial, non-profit or government organization with 1,000 or more employees. Respondents hold various cybersecurity roles, including chief information officer (CIO), chief information security officer (CISO), VP, director or manager of IT security, and IT security administrator.

Research participants were drawn from 10 countries and more than 15 industries. With a sample size of 1,000 participants, the global survey margin of error (at a standard 95% confidence level) is 3%. All results pertaining to individual countries and industries should be viewed as anecdotal, as their sample sizes are much smaller. Proofpoint recommends making actionable decisions based on global data only.

Proofpoint platform data

Data was sourced from deployments of the Proofpoint human-centric security platform between April 2024 and May 2025.

About CyberEdge Group

Founded in 2012, CyberEdge Group is the premier research, marketing, and publishing firm dedicated exclusively to serving the cybersecurity vendor community. As the producer of the distinguished Cyberthreat Defense Report (CDR) and numerous other award-winning research studies, CyberEdge has earned recognition from top-tier business and technology outlets, including The Wall Street Journal, Forbes, Fortune, USA Today, NBC News, ABC News, SC Media, Dark Reading, CISO Magazine, and Security Buzz.

Renowned for its depth of cybersecurity expertise and commitment to excellence, CyberEdge delivers world-class market research, survey analyses, analyst reports, white papers, and custom books and eBooks tailored to the cybersecurity industry. Its unmatched combination of subject-matter knowledge and comprehensive service offerings continues to set the gold standard for quality and insight.

To learn more, visit www.cyberedgegroup.com.

proofpoint.

Proofpoint, Inc. is a global leader in human- and agent-centric cybersecurity, securing how people, data and AI agents connect across email, cloud and collaboration tools. Proofpoint is a trusted partner to over 80 of the Fortune 100, over 10,000 large enterprises, and millions of smaller organisations in stopping threats, preventing data loss, and building resilience across people and AI workflows. Proofpoint's collaboration and data security platform helps organisations of all sizes protect and empower their people while embracing AI securely and confidently. Learn more at www.proofpoint.com.

Connect with Proofpoint: LinkedIn

DISCOVER THE PROOFPOINT PLATFORM →