proofpoint®

E-BOOK

Transform Data Security

Rethink your data protection strategy to align with how business operates today

Preventing data loss in an era of transformative change

A software company accuses a competitor of planting a spy to steal thousands of pieces of proprietary and confidential information.[1] A major bank mistakenly sends a customer hundreds of pages of other clients' sensitive investment information.[2] An employee at a global electronics company pastes proprietary source code into ChatGPT for debugging.[3] A healthcare organization accidentally emails the wrong file, exposing the personal data of 14,000 employees.[4]

Each of these real-world incidents is different. But they share a common outcome: data loss with serious consequences. These include financial damage, reputational harm, regulatory exposure and erosion of customer trust.

Data loss isn't new. It can stem from both external attacks and insider behavior, whether careless or malicious. But today's enterprise environments make preventing it far more challenging. With corporate information spread across cloud platforms and on-premises systems, workers are putting data at risk in increasingly varied ways. Meanwhile, security teams are stretched thin, expected to do more with fewer resources.

Adding to this challenge, compliance demands have grown more complex and the penalties more severe.

Every organization is unique. But one thing is always true: data doesn't leak itself. People move, misuse and leak data.

This e-book:

- Explores the challenges of securing data in today's dynamic work and IT environments
- Explains why traditional data security solutions are no longer sufficient
- Offers a strategic roadmap for transforming enterprise data security to meet the needs of the modern workplace

71%

of respondents stated that a careless user was the cause of data loss.[5]

20%

of respondents stated that a malicious employee or contractor was the cause of data loss.[6]

1. SFGate. "San Francisco tech company claims it caught rival in 'brazen' spying attack." March 2025.
2. Financial Times. "Lloyd's blames 'human error' after sending customers other clients' investment data." March 2025.
3. Mashable. "Whoops, Samsung Workers Accidentally Leaked Trade Secrets via ChatGPT." April 2023.
4. Infosecurity Magazine. "Data Leak Hits Thousands of NHS Workers." February 2023.
5. Proofpoint. *Data Loss Landscape Report*. 2024.
6. Ibid.

SECTION 1

Data security and the modern organization

Data risk today is driven by people working in increasingly complex environments. Insider threats, whether from careless, malicious or compromised users, now cause over a third of breaches.[7]

At the same time, data sprawl across clouds, software as a service (SaaS) applications and large language models (LLMs) creates blind spots for security teams. For example, security researchers found over 12,000 API keys and passwords in public datasets used for LLM training.[8]

In addition to "shadow" data stored outside the visibility of IT teams, abandoned enterprise data also drives up risk and cost. For example, data stored in multi-cloud environments can become redundant and backups or snapshots abandoned.

This unused data drives up cloud storage and backup costs. It also increases an organization's attack surface without delivering value.

Data risk is a people problem

Workers increasingly put data at risk—and in more dangerous ways. Sometimes, this risk is accidental, such as a recruiter emailing sensitive candidate data to the wrong contact or a developer pasting proprietary code into a public generative AI (GenAI) tool.

Other times, the threat is intentional, such as a departing employee exfiltrating intellectual property (IP). Increasingly, data exposure comes from compromised insiders whose stolen credentials are used by attackers to access sensitive systems.

Careless users may make an honest mistake or try to take a shortcut to do their jobs.

Malicious users can intentionally exfiltrate data for personal gain.

Compromised users may have their accounts taken over and misused by an outside cyberattacker.

$17.4M

is the total average cost of an insider data breach.[9]

85%

of organizations experienced one or more data loss incidents in the past year.[10]

7. Verizon. *2024 Verizon Data Breach Investigations Report*. May 2024.
8. The Hacker News. "12,000+ API Keys and Passwords Found in Public Datasets Used for LLM Training." February 2025.
9. Ponemon Institute. *2025 Cost of Insider Threats Global Report*. February 2025.
10. Proofpoint. *Data Loss Landscape Report.* 2024.

1 in 3 breaches involved shadow data and drove a 16% higher average cost per incident.

Data is everywhere and anywhere

Data sprawl has become a major challenge. Today's knowledge workers store data across public clouds, file-share services and on-premises databases. This fragmentation creates blind spots for IT and security teams.

Most workers create, copy or share sensitive data without malicious intent. However, when this data is stored in an unsanctioned SaaS application, it can be an accident waiting to happen. The 2024 Ponemon Institute and IBM *Cost of a Data Breach* report found that one in three breaches involved shadow data. This drove a 16% higher average cost per incident.[11]

Compounding these issues, cloud data stores often become abandoned. Much of this forgotten data might be unknown to IT, making it challenging to protect.

11. Ponemon Institute and IBM. *2024 Cost of a Data Breach Report.* July 2024.

More demand for AI in the enterprise

Use of AI—particularly generative AI (GenAI)—is surging in modern enterprises. In a 2025 McKinsey report, 78% of organizations stated that they already use AI in at least one business function.[12]

Strong data protection is critical for preventing leakage of sensitive information through GenAI. But CISOs can't simply block GenAI tools: they need to balance security with productivity and innovation. Unfortunately, legacy DLP solutions aren't effective at managing this delicate balance. According to Harmonic, 8.5% of GenAI prompts include sensitive data.[13] And according to the 2025 McKinsey report, 73% of organizations experienced at least one negative consequence from GenAI.[14]

Security teams must discover and classify sensitive data so that it isn't exposed through public GenAI tools, enterprise AI tools such as Microsoft Copilot, or through custom LLMs. They need visibility of who uses GenAI services, the intent behind requests and whether sensitive data is being shared. They must also ensure that their own custom LLMs or retrieval-augmented generation (RAG) applications protect sensitive or proprietary data.

8.5%

of GenAI prompts include sensitive data.[13]

73%

of organizations experienced at least one negative consequence from GenAI.[14]

Security teams struggle to keep up

While enterprise environments are becoming more distributed and dynamic, security and IT teams are being asked to do more—with less. Economic pressures can cause organizations to reduce budgets and freeze hiring, even as volumes of enterprise data grow and its sprawl across on-premises and cloud environments becomes an increasing data security challenge.

Fragmented IT stacks add to this pressure. Many security programs rely on siloed data security tools that don't integrate well. This creates visibility gaps, operational overhead and longer response times.

Regulatory complexity is also growing. Compliance frameworks keep evolving and vary by region, industry, and data type. Organizations must continuously adapt to business demands such as securing data during a restructuring, often with limited resources.

12. McKinsey. *The state of AI: How organizations are rewiring to capture value.* March 2025.
13. Harmonic. *From Payrolls to Patents: The Spectrum of Data Leaked into GenAI.* 2024.
14. McKinsey. *The state of AI: How organizations are rewiring to capture value.* March 2025.

SECTION 2

A new mandate for data security teams

The ways organizations create, store, use and manage data have changed. So have data security priorities. To understand how to transform your organization's data security strategy, you must first recognize how legacy solutions fail to meet today's biggest data security challenges.

Today's data security priorities

Modern organizations increasingly identify three top priorities for data security: reducing risk, lowering operational costs and enabling business agility.

Reducing security risk

To reduce risk, enterprises must have clear visibility of where sensitive data lives, who or what can access it, how it's being used and how it might be exposed. This applies across email, cloud endpoints, and data stores, whether cloud, SaaS, or on-premises. By combining accurate data discovery with detection of risky behavior, security teams can understand both the content of, and the intent behind, user actions. This enables faster, more accurate responses—often finding data exposures before they can be exploited and stopping data loss before it happens.

Lowering operational costs

Lowering costs means speeding up deployments, reducing tool sprawl and making data security operations more efficient. A streamlined solution must simplify the complex and fragmented data security challenges modern organizations face. These include locating their data, classifying it correctly, controlling who can access it, and monitoring how people interact with it across all channels—from endpoints and email to the web and the cloud. With AI-driven protection that doesn't require manual data classification, businesses can launch strong security programs in weeks, not years.

Enabling business agility

Today's organizations must stay secure while moving quickly—whether expanding, restructuring or innovating through the adoption of new technologies such as GenAI.

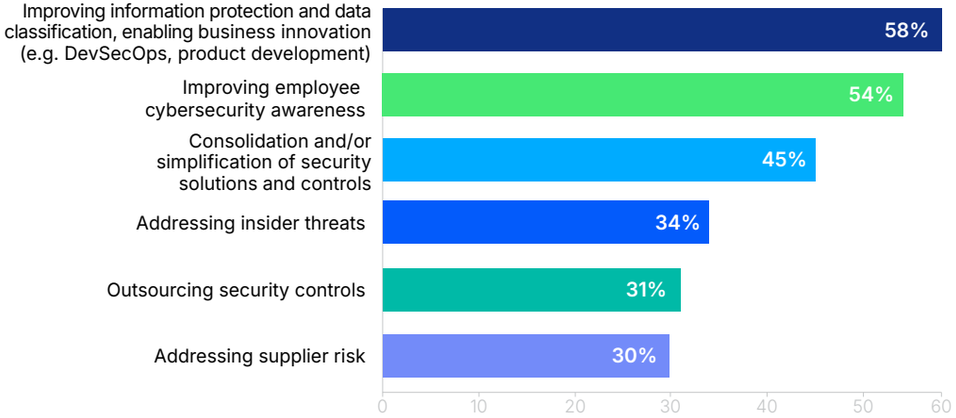
The broken promises of legacy data security tools

As part of their information protection and compliance programs, many organizations rely on a patchwork of independent data security, insider risk, and data access governance tools. This fragmented approach is costly, complex to manage and often disruptive to end users.

Worse, these tools fail to deliver on their already-limited promises. As a result, 72% of organizations use multiple data security tools to achieve their goals.[15]

Considering this, it's not surprising that 58% of CISOs are focused on improving information protection and data classification:

What are the top priorities for your organization's IT security department over the next two years? *(Pick up to three.)*



Source: Proofpoint. *Voice of the CISO Report.* 2024.

Legacy data security tools miss important context

Legacy data security tools analyze data events in isolation, ignoring who's involved or why events happened. But data security must go beyond just data awareness. To accurately detect risks and enable fast responses, it must also be context-aware.

There are some key reasons why legacy information protection tools struggle in this area:

- Limited visibility of who can access and interact with sensitive data in the cloud, by email or on endpoints
- Lack of user monitoring and insider threat detection
- Inability to identify user activities that don't trigger alerts, but provide important context for threat detection

72% of organizations use multiple data security tools to achieve their goals.[15]

15. Cloud Security Alliance. *Data Loss Prevention and Data Security Survey Report*. 2023.

Data loss detection policies are hard to write, easy to evade

Traditional data security tools are built for regulated data that's easy to detect and hard for users to alter. To avoid false positives, policies must be precise and granular. This level of policy detail is manageable when the data is stored in few static locations.

However, sensitive business information, regulated data and IP now appear almost anywhere, in all kinds of files and documents. This makes traditional approaches impractical and ineffective. The result is policies that are hard to write and easy to evade.

Sensitive data—particularly in unstructured formats—can be hard to recognize and label. And users can easily change those labels, either for legitimate reasons, by accident or maliciously.

An imbalance between security and productivity

Traditional information protection tools often use strict blocking policies that don't account for modern, flexible workstyles.

Blocking access to personal email, cloud tools or the public internet can frustrate users and disrupt their work.

Take GenAI tools, for example. Legacy solutions can't stop sensitive data from being included in GenAI prompts, instead relying on broadly blocking such tools. But this only drives usage underground, increasing risk and reducing visibility.

Traditional data discovery and classification take too long

Legacy DLP delivers slow, incomplete data discovery. It relies on rigid patterns found in structured data, such as personally identifiable information (PII) or protected health information (PHI). But this approach breaks down with unstructured data, such as IP.

Worse, legacy tools need full data classification to be effective—a process that can take months or years. And most can't use existing classification schemes, such as Microsoft Information Protection (MIP) labels, without costly add-on services.

Too many false positives lead to alert fatigue

Incident responders are hindered by the inaccuracy of traditional tools. To stop potential breaches, they must act fast. But they often have a high volume of data security alerts to triage, many of them false positives.

In the Proofpoint 2024 *State of the Phish* report, 80% of respondents said that one-fifth of their cybersecurity alerts were false positives.[16] More alarming, 55% said that their teams missed critical alerts because they couldn't prioritize alerts effectively.[17]

80% of respondents say that one-fifth of their cybersecurity alerts are false positives.[16]

55% of respondents say that their teams missed critical alerts because they couldn't prioritize alerts effectively.[17]

16. Proofpoint. *State of the Phish Report*. 2024.
17. Ibid.

SECTION 3

A unique approach to data security

Transforming data security starts with people. Your new approach must adapt to risk, understand intent and provide visibility across every user, channel, and data store.

For modern enterprises, an optimal approach to data security is:

- **Human-centric:** Detects risky behavior early by monitoring user intent and content
- **Adaptive:** Uses AI to automatically adjust protections based on real-time risk
- **Comprehensive:** Provides unified visibility of all users—careless, malicious or compromised. Protects every channel (email, endpoint, cloud) and data store (from on-premises to cloud).

This transformative approach combines behavioral analytics and automation, helping teams reduce data loss risk, simplify operations and support business agility. According to Gartner, organizations incorporating intent detection and real-time remediation capabilities into their data security programs will see a one-third reduction in insider risks by 2027.[18]

Human-centric data security understands content and behavior

Human-centric data security requires deeply understanding your data. This means knowing where it resides, how sensitive it is and who can access it—across every digital channel and data store. It also means going beyond content alone to understand user behavior and intent. With these extra insights, your organization can identify whether users are careless, malicious or compromised.

Human-centric data security rests on two core pillars:

- **Content awareness:** Accurately discovering and classifying sensitive data across all digital channels, using techniques such as AI-based classification, labeling, exact data matching and proximity analysis
- **User behavior awareness:** Interpreting user activity across channels and data stores, enabling detection of risky actions in context

18. Gartner. *Market Guide for Data Loss Prevention.* April 2025.

To illustrate this, the following example describes a human-centric approach to data security:

A financial services firm detects an employee accessing thousands of customer records late at night, well outside their normal activity and working hours. Rather than automatically blocking the action, the system uses behavioral context to trigger a real-time, in-app prompt asking whether the activity is work-related. If confirmed, the explanation is logged. If not, security is alerted for review.

In this situation, the employee unintentionally triggered a bulk export during testing. Using a context-aware prompt, the potential incident was resolved without disruption, embarrassment, or halting legitimate work. The user was treated as a trusted partner in security and data stayed protected.

Adaptive data security learns and adjusts controls in real time

An adaptive data security platform uses AI-based classification and detection to continuously learn and understand data sensitivity and data interactions, without reconfiguring rules. AI-powered classification provides deeper visibility of sensitive data and automatically labels content to enforce encryption, access restrictions and editing rights.

In addition, AI-driven solutions continuously adjust to changing risk levels—both for users and for data. By understanding who is most at risk and how they interact with sensitive content, these solutions apply the right controls in real time. They analyze user behavior in context to reduce false positives. They also tailor policies based on shifting roles and access.

This adaptive approach enables security teams to scale with the business—without compromising protection.

Comprehensive data security is unified

A comprehensive data security solution unifies visibility and control across all exfiltration channels, data sources and user types. It tracks and interprets data movement and user behavior across your organization's whole environment. And because it adapts as human and data risks change, it comprehensively protects against data exfiltration, data exposures and insider threats.

From sales to HR, finance to engineering, a unified data security platform gives you the complete visibility and adaptive control needed to protect data used by everyone in your organization—wherever they do their work.

Advantages of human-centric data security

A human-centric, adaptive and comprehensive approach to data security provides significant benefits. By unifying data protection, using AI for faster, more accurate classification and detection, and applying context-aware responses, organizations can better prevent data exfiltration, exposure, and insider threats. The following table shows key capabilities and features that deliver these benefits for security, compliance, and operational efficiency.

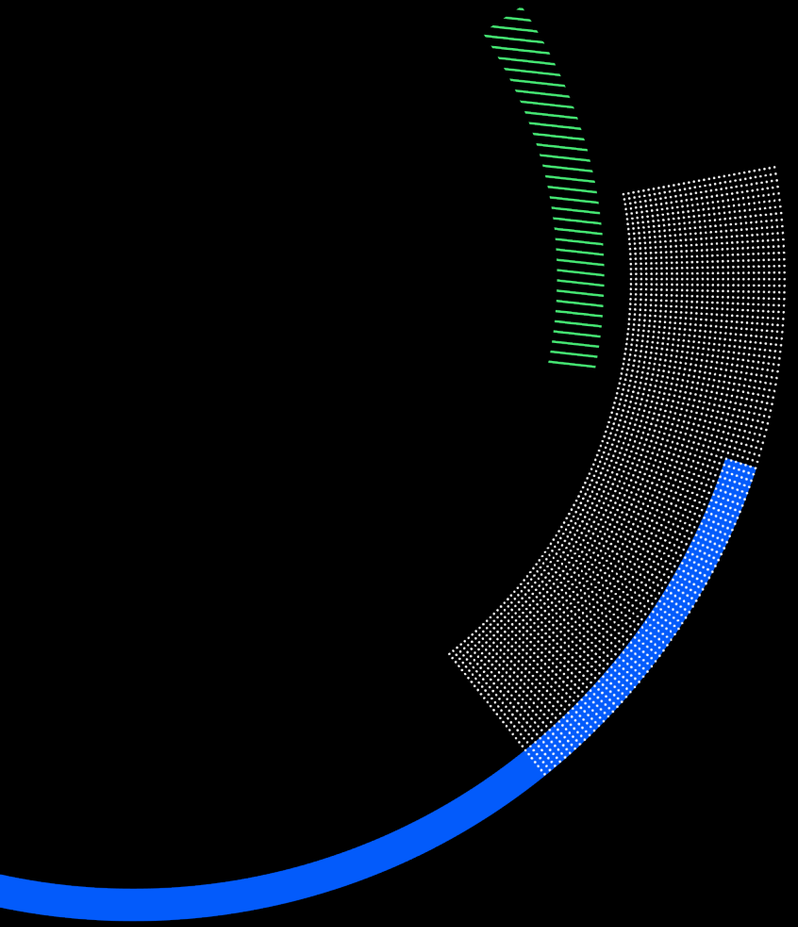
CAPABILITIES	FEATURES
 **Unified data security**	• **Single architecture** integrating DLP, data security posture management (DSPM) and insider threat management (ITM). • **Unified policy engine and classification** across all channels and data stores. Consistent security and higher alert fidelity.
 **Intent-aware incident response**	• **Intent-based controls** based on detection of risky behaviors (careless, malicious, compromised). • **Full data lineage** that tracks sensitive data from creation through every interaction. Critical context for incident response. • **Comprehensive protection** that locates sensitive data, governs access, prevents loss and continuously monitors risky actions. • **Context-aware responses** including user nudges, justification prompts and real-time notifications. More refined control than simple allow or block actions.
 **AI-driven data security**	• **Auto-learning, AI-powered classification** of data and documents that are new or unique to your organization. Full visibility and control of data sprawl. • **Anomaly detection** that identifies unusual user activities, such as downloading large volumes of data from cloud folders or changing file permissions. • **Adaptive enforcement** that updates policies and takes appropriate action when a user's risk level changes. Reduced manual effort. • **Automated remediation** that removes excessive privileges and addresses cloud misconfigurations. Less manual effort and exposure risk.
 **Compliance and privacy**	• **Compliance framework** that automatically maps data configurations against legal and regulatory requirements to protect sensitive data. • **Privacy by design** with built-in mechanisms that help ensure compliance with global and local data privacy regulations.
 **Operational efficiency**	• **Quick, seamless deployment** that accelerates time to value. • **Scalable architecture** that grows as organizations evolve. • **Elegant user experience** that protects sensitive information without affecting end users, analysts and administrators.

Conclusion: It's time to transform data security

Transforming data security requires a shift towards a human-centric, adaptive and comprehensive approach. One that understands behavior and intent, dynamically adjusts to changing risks, and delivers full visibility across data, users, and threats—without hindering productivity. It's not just about preventing data loss: it's about empowering security teams to act faster, protect more and support innovation with confidence.

If your data security strategy still relies on yesterday's tools, it's time to rethink what's possible. Choose a platform that aligns with how people work today and how your organization will grow tomorrow.

Learn more about how Proofpoint can help you transform your data security. Visit https://www.proofpoint.com/us/products/data-loss-prevention.



proofpoint.

Proofpoint, Inc. is a leading cybersecurity and compliance company that protects organizations' greatest assets and biggest risks: their people. With an integrated suite of cloud-based solutions, Proofpoint helps companies around the world stop targeted threats, safeguard their data, and make their users more resilient against cyber attacks. Leading organizations of all sizes, including 85% of the Fortune 100, rely on Proofpoint for people-centric security and compliance solutions that mitigate their most critical risks across email, the cloud, social media, and the web. More information is available at www.proofpoint.com

Connect with Proofpoint: LinkedIn

DISCOVER THE PROOFPOINT PLATFORM →