proofpoint®

E-BOOK

Transforming Data Security

Rethinking your data protection strategy to align with how people work today

Preventing data loss in an era of transformative change

A professional services firm accuses a competitor of planting a spy to steal thousands of confidential client records.[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 consequences 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. But today's enterprise environments make the challenge far more complex. With corporate information dispersed 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 are growing more complex and the penalties for failure are more severe than ever.

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

Reducing the risk of sensitive data loss requires more than incremental improvements. It demands a new approach.

This e-book explores the challenges of securing data in today's dynamic work and IT environments. It explains why traditional data loss prevention (DLP) solutions are no longer sufficient. And it 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]

58%

of organizations had compromised cloud accounts.[7]

30%

of tenants experienced post-compromise data exfiltration or manipulation.[8]

1. SFGate. "San Francisco tech company claims it caught rival in 'brazen' spying attack." March 17, 2025.
2. Financial Times. "Lloyd's blames 'human error' after sending customers other clients' investment data." March 17, 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.
7. Ibid.
8. Ibid.

SECTION 1

Data security and the modern organization

Data risk today is driven by people, complexity and consequence. Insider threats — whether careless, malicious, or through compromised accounts — now cause over a third of breaches.[9] At the same time, data sprawl across clouds, software-as-a-service (SaaS) apps and large language models (LLMs) creates blind spots. Shadow data in these locations drives up costs and compliance risk. Security teams face these challenges with limited resources, fragmented tools and mounting regulatory demands. The result: data breaches are more frequent, more expensive and more damaging to businesses than ever.

Data risk is a people problem

Workers are increasingly putting data at risk — and in more dangerous ways. According to the *2024 Verizon Data Breach Investigations* report, 35% of breaches involved internal actors. This value is a sharp increase from 20% the year before.

Sometimes, the cause is accidental. For example, a recruiter emailing sensitive candidate data to the wrong contact or a developer pasting proprietary code into a generative AI (GenAI) tool.

Other times, the threat is intentional, such as a departing employee exfiltrating intellectual property to benefit a competitor. And increasingly, it stems from compromised insiders, where attackers use stolen credentials to act as trusted users and access sensitive systems undetected.

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

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

Malicious users can intentionally exfiltrate data for personal gain.

43%

of employees whose jobs can be done remotely are working from home all or most of the time.[10]

$17.4M

is the total average cost of an insider data breach.[11]

85%

of organizations experienced one or more data loss incidents in the past year.[12]

9. Verizon. *2024 Verizon Data Breach Investigations Report*. May 2024.
10. Flex Index. *The Flex Report Q4 2024.* December 2024.
11. Ponemon. *2025 Cost of Insider Threats Global Report*. February 2025.
12. Proofpoint. *Data Loss Landscape Report.* 2024.

Data is everywhere and anywhere

Data sprawl has become a critical challenge. Organizations have embraced the cloud. Developers and knowledge workers store data across a complex mix of environments. These include public clouds such as Amazon Web Services and Google Cloud Platform, file-share services such as Microsoft OneDrive, LLMs and on-premises databases. This fragmentation creates blind spots for IT and security teams. As a result, it's harder to locate sensitive data, detect exposure and show compliance.

Workers often create, copy or store sensitive data without malicious intent. However, when this data is outside the visibility or control of an IT team, 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 this kind of shadow data.This drove a 16% higher average cost per incident.[13]

IT teams struggle to keep up

Threat landscapes are expanding and enterprise environments are increasingly distributed and dynamic. At the same time, security and IT teams are being asked to do more — with less. Budget constraints, hiring freezes and macroeconomic pressures often force organizations to scale back resources. This is happening even as the volume and sophistication of cyber risks continue to rise.

Work and personal lives have blurred

Personal devices are used for work; and employer-provided devices are used by the family and for leisure. This trend complicates even the most well-managed data security and data privacy efforts.

Alongside remote work, organizations are embracing the cloud. Users rely on software-as-a-service (SaaS) platforms, cloud storage, collaboration tools, chat and videoconferencing.

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

A fragmented technology stack adds to this pressure. Many security programs rely on a patchwork of siloed tools that don't integrate well. This creates visibility gaps, greater operational overhead and longer response times. Without a unified view of risk, security teams have to stitch together insights from disparate systems. This slows down decision-making and increases the chances of missed threats.

Talent shortages add a further challenge. Skilled cybersecurity professionals are in high demand but in short supply. As a result, even high-performing teams might be overwhelmed and reactive, chasing alerts instead of executing a proactive, strategic security roadmap.

Regulatory complexity also continues to increase. Compliance frameworks evolve rapidly and vary by region, industry and data type. These include General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), Health Insurance Portability and Accountability Act (HIPAA), Payment Card Industry Data Security Standard (PCI DSS) and new AI governance mandates. Organizations must continuously adapt policies, processes and controls — often without dedicated compliance resources.

Data breaches have real business consequences

Data loss is no longer just an IT concern — it's a major business risk. In 2024, the average data breach cost rose 10% from the previous year to $4.88 million, with malicious insider incidents averaging $4.99 million.[14] These rising costs reflect technical recovery, operational disruption and long-term financial impact.

The reputational fallout is just as damaging. In 2024, 85% of organizations experienced data loss in the past year.[15] Of those, more than 90% suffered consequences — most commonly revenue loss (57%) and reputational damage (39%).[16] In a trust-driven economy, these effects can be lasting.

Regulatory consequences continue to grow as well. Breaches now routinely lead to fines, audits and legal exposure under evolving frameworks such as GDPR and CCPA. With sensitive data scattered across environments and in constant motion, organizations must rethink how they prevent loss before it becomes a business crisis.

14. Ponemon. *2024 Ponemon Institute and IBM Cost of a Data Breach Report*, July 2024.
15. Proofpoint. *Data Loss Landscape Report.* 2024
16. Ibid.

SECTION 2

A new mandate for data security teams

The ways we create, store and use data have changed. The categories of users who are given access to that data have changed. The risks and potential impacts have changed.

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

Reducing security risk requires a proactive approach. This means going beyond monitoring data movement to also understanding user behavior and intent. By detecting risky actions across email, cloud and endpoints, organizations can prevent data loss before it happens. Enhanced visibility of where data lives and how it's used also helps security teams manage sprawl and meet compliance obligations.

Lowering operational costs is another pressing need. Today's security teams are burdened by too many siloed tools, leading to fragmented investigations and limited insights. Consolidating platforms and streamlining DLP and insider risk programs not only reduces overhead but improves efficiency. Smarter data classification also plays a role — eliminating redundant storage, reducing exposure and strengthening overall security posture.

Enabling business agility means staying secure while moving fast. Whether undergoing mergers, restructuring, or adopting transformative technologies such as GenAI, organizations must be ready to respond to new risks without slowing innovation.

Data security: the goal

The main goal of a data security program is to stop people moving sensitive or critical data out of the organization in ways that are risky or violate policy. Increasingly, data security programs are critical to complying with industry and regional privacy regulations concerning personal data.

Data security technologies have a dual challenge. On the one hand, they must ensure that people access and use sensitive or critical data appropriately. On the other, they should not block business transactions or hinder user productivity.

Keeping the trust

Here are a few examples of sensitive or confidential data that data security is designed to protect:

- Personally identifiable information (PII) for employees and customers
- Personal health information (PHI) for employees and customers
- Personal financial and banking informationfor employees and customers
- Trade secrets
- Intellectual property
- Customers lists
- Vendor information
- Material information
- Other sensitive business information

Data security: the reality

Most organizations have adopted traditional DLP technologies to stay compliant. But these tools are often expensive. They're cumbersome to maintain. They get in users' way. And, worst of all, they fail to deliver on their already-limited promises.

Organizations need a better way to protect their sensitive data. Large organizations can't use just one DLP tool. Instead, 72% of them cobble together two or more tools as a part of their data security strategy.[17]

And organizations aren't just using more tools to combat insider threats. They're also spending more than ever before.

Since 2016, the average cost of managing insider threats has soared more than 85%. Investigation costs for these threats have jumped 17% in only two years.[18]

A data-centric focus misses important context

Data doesn't move itself. People lose and misuse data. Protecting data requires knowing the context of what data your people have access to, what they're doing with it and how they're being targeted by cyberattackers that seek access. In other words, data-aware DLP is not enough — it must also be people-aware and threat-aware.

Traditional DLP typically overlooks context for these reasons:

- It has limited or no visibility into who is interacting with and moving sensitive data. This includes touchpoints such as the cloud, the web, email, print, USB devices and endpoints.
- It lacks insider threat detection and response.
- It cannot join the dots between activities that don't rise to the level of an alert but which are critical in context.
- It isn't integrated with a threat protection platform or real-time threat intelligence.

$7B

total projected DLP spending by 2028.[19]

81 days

to resolve insider threats.[20]

96%

of organizations are targeted by cloud attacks.[21]

70%

of incidents relate to user carelessness.[22]

17. Cloud Security Alliance. *Data Loss Prevention and Data Security Survey Report*. 2023.
18. Ponemon. 2025 *Cost of Insider Threats Global Report*. February 2025.
19. The Radicati Group. "Data Loss Prevention - Market Quadrant 2024." March 2024.
20. Ponemon. *2025 Cost of Insider Threats Global Report*. February 2025.
21. Assaf Friedman and Itir Clarke (Proofpoint). "How Attackers Use Compromised Accounts to Create and Distribute Malicious OAuth Apps." May 2021.
22. Proofpoint. *2024 Data Landscape Report*. April 2024.
23. Proofpoint. *2023 State of the Phish Report*. February 2023.

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

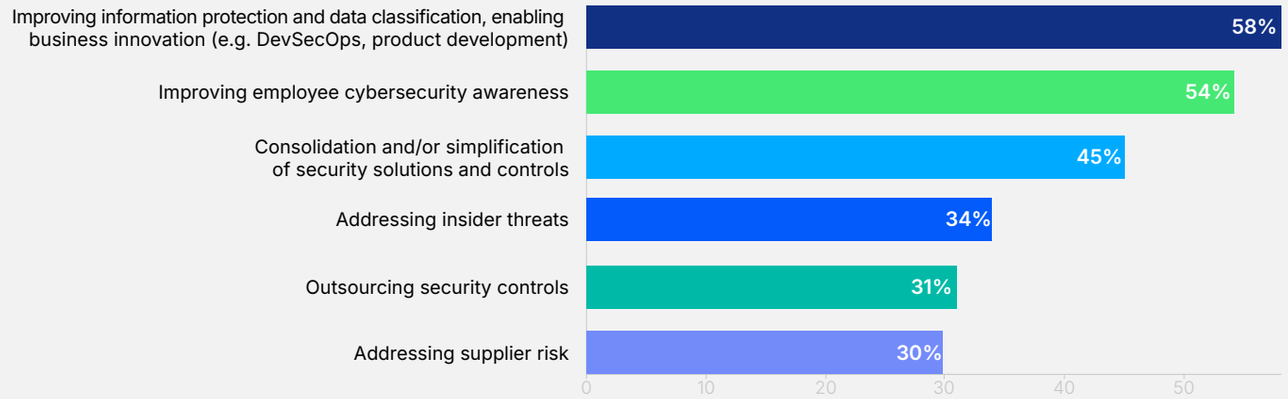
Traditional DLP tools were built for regulated data that's easy for tools to detect and hard for users to alter. To avoid false positives, they require policies that are precise and granular. A typical policy might include traits such as data identifiers, application name, data exfiltration channel and so on.

This level of policy detail was manageable when the data was stored in databases, on servers and in a few static locations.

Today, sensitive business information, regulated data and intellectual property can appear almost anywhere, in all kinds of files and documents. This shift has made DLP policies much harder to write and far easier to evade.

Sensitive data and files are much harder to recognize and label with simple data identifiers. And users can easily change them — whether for legitimate business reasons, by accident or through malicious actions.

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.

Countering such actions requires writing and maintaining a complex list of overlapping DLP policies. Any telemetry on data movement outside of a detected alert is stored in log files that are, at best, hard to access and analyze. And no user behavior or threat telemetry is recorded at all. If security analysts want to piece together these events into a narrative sequence for context, they have to do it manually, or by using several disparate tools.

It's no wonder that most security teams have such little visibility of data movement across their organizations.

Limited prevention of careless users' mistakes

Clearly, these narrowly defined policies limit how well traditional DLP tools can protect data. They also fail to address accidental or careless user behavior that exposes sensitive data. To understand why this is such a problem, consider an all-too-common example of user carelessness: misdirected email.

It's easy for hurried users to make mistakes when they're adding recipients to their emails because most webmail and native email clients now offer address autofill. According to 2023 data from Tessian (now a Proofpoint company), about one-third of users send about two emails to the wrong person annually. That means a business of 5,000 employees can expect to deal with around 3,400 misdirected emails per year.

Or consider the issue of hard-to-decipher controls. These can lead users to put data at risk by sharing it publicly or with GenAI tools, rather than just within their organizations. Data can be compromised simply because traditional DLP tools weren't built for malicious or risky user actions.

Controls burden users, strain systems

Besides resource-heavy endpoint agents, other aspects of traditional DLP tools can strain system resources and burden users with controls that get in the way of real work.

The more organizations need from their DLP tools, the less they get. They might try to set up complex and overlapping detection and prevention policies that require deeper content inspection only to see slower performance as a result.

Why users hate legacy DLP tools

Traditional DLP tools can also be overzealous in their blocking polices. With remote and hybrid work styles as the norm, security teams can't anticipate all the ways people choose to work. When organizations rely on blocking policies, they stop their people from using the public internet, accessing personal email at work or working with new cloud-based tools. Take GenAI tools as an example. Legacy DLP tools can't prevent data loss via prompt submissions. Blunt security practices such as web filtering for GenAI sites only drive employees underground. This stops security teams from understanding how the technology is used and increases overall risk.

Ultimately, security teams can't block employees from taking risky actions and putting themselves and the organization at risk. Even the best technical defenses can be undermined if users don't follow basic best cybersecurity practices. In a recent Proofpoint survey, 71% of users said that they took a risky action, and 96% were aware of the dangers but did so anyway.

96%

of users know when they are taking a risky action but do it anyway.[23]

24. Proofpoint. *2024 State of the Phish Report.* 2024.

When blocking policies become stumbling blocks

Some DLP systems may block legitimate processes because those processes aren't listed within DLP policies and approved lists. This is a major issue for modern businesses, which rely on friction-free digital transactions and communication. When DLP policies mistakenly block a legitimate transaction or message, security teams are left dealing with users' complaints.

In the worst cases, the endpoint DLP solution clashes with another endpoint security tool or causes system crashes. These have real bottom-line impacts on employees, customers and partners.

Traditional data discovery and classification take too long

With traditional DLP solutions, discovering and classifying data can take months. Instead of focusing on how users are moving data in real time, in potentially risky ways, traditional DLP focuses on what happened to data in the past.

Without AI-assisted scanning tools, the data discovery process is slow. It also misses data stores that span an organization's entire environment. Searches usually run only outside business hours to avoid the heavy toll they take on system performance and productivity.

Another problem is that most DLP tools cannot ingest prior data classification efforts — including Microsoft Information Protection — without high-priced professional services. With each new program, organizations must reclassify their data in the new tool.

Deployment and maintenance are complex and expensive

Deploying traditional DLP tools — especially on premises — is often complex and expensive. Installing and integrating servers, applications, databases and other infrastructure can take months. Time to value can be even longer.

Endpoint DLP products can be just as problematic. Many use kernel-mode endpoint agents. These agents intercept every OS-level transaction and thus are heavy on the endpoint. They can slow down users' work, interfere with applications and even crash the device. In many cases, these issues are so severe that they crop up in early testing stages.

Once installed, many traditional DLP solutions require setting up and maintaining complex rules and policies — a major investment of time and money.

80% of respondents say that one-fifth of their cybersecurity alerts are false positives.[24]

False positives lead to alert fatigue

Incident responders are also frustrated by the lack of accuracy of traditional DLP solutions. To halt a potential breach in progress, they need to act fast. But they often have difficulty dealing with the sheer volume of DLP alerts, many of them false positives.

In one survey, 80% of respondents said that one-fifth of their cybersecurity alerts are false positives. What's more alarming, more than half (55%) said that their teams have missed critical alerts because they don't prioritize alerts effectively.[24]

25.

SECTION 3

The modern data security difference

Modern data security starts with people. It adapts to risk, understands intent and delivers full visibility across every channel and user.

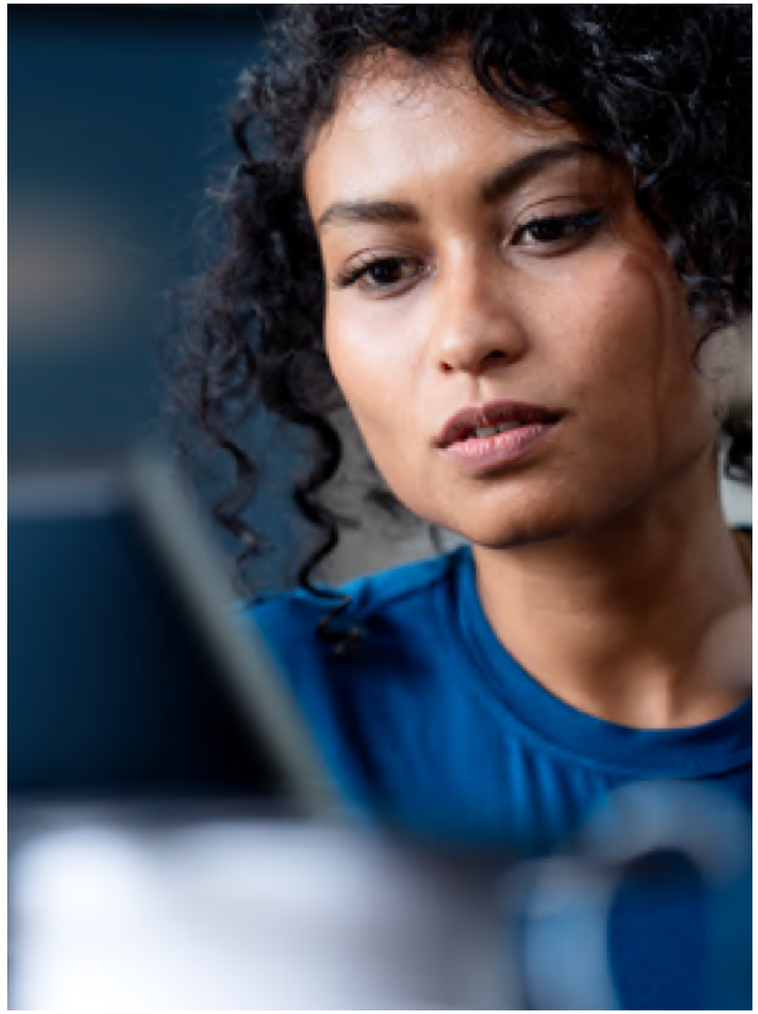
A modern approach to data security must be human-centric, adaptive and comprehensive. That means detecting risky behavior early by monitoring intent — not just content — across all channels. It means using AI to dynamically adjust controls based on real-time risk. And it means unifying visibility of all users, whether they are careless, compromised, or malicious.

Modern data security meets those needs by combining behavioral analytics, automated response and full coverage across endpoints, cloud and email. This means that data protection teams can reduce risk of data loss, streamline operations and support business agility.

Advantages of a modern approach

Here are just a few of the benefits of modern data security over traditional DLP:

- **Assessments of human-driven risks and detection of** behavioral anomalies
- **Context to discern** malicious, compromised and careless insiders
- **Consistent and adaptive controls** based on human-risk for data at rest and in motion
- **Cloud-scale data discovery and classification** and posture management
- **Unified content intelligence** for data at rest and in motion that enables organizations to discover, classify and protect unique sensitive content
- **Large language model (LLM) classifiers** that detect valuable unstructured data such as human resources (HR) documents
- **Ability to detect sensitive data in images**



- **Automated remediation actions** for revoking access to compromised identities, adjusting file-sharing permissions and quarantining affected files
- **Better protection for sensitive data** and intellectual property, including acceptable use policies for GenAI tools
- **A unified console**, streamlining analyst workflows and accelerating investigations
- **Enhanced data lineage capabilities** that trace the history of a data over time as it traverses various channels, based on user behavior, intent and content
- **Context-aware controls** that go beyond allow and block activities and include user nudges, justifications and notifications
- **Data controls** for managed and unmanaged devices
- **Privacy by design** to keep pace with increasing data privacy regulations all over the world
- **Extensibility** to work in concert with the broader security ecosystem — without the need for significant engineering effort
- **Fast and easy deployment**, accelerating time to value
- **Scalability** as organizations grow and change

Modern data security is human-centric

Human-centric data security means going beyond content inspection to also understand user behavior and intent. A modern strategy must correlate user actions with data movement across files, applications and endpoints. This includes identifying when users have excessive or inappropriate access to sensitive information. The extra behavioral context enables security teams to anticipate risks, detect threats earlier and accelerate investigations across all digital channels.

Human-centric data security rests on three core pillars:

- **Content awareness** to accurately discover and classify sensitive data across all digital channels, using techniques such as labeling, exact data matching and proximity analysis
- **User behavior awareness** to interpret activity across cloud, endpoints and email channels — enabling detection of risky actions in context
- **External threat awareness** to correlate compromised users and phishing campaigns with insider risks

Modern data security is adaptive

People and data carry varying levels of risk. Some users are prime targets for attackers due to their roles or access. Others are simply more prone to human error. Similarly, not all data is equally sensitive or equally exposed. Activity around critical data fluctuates constantly, and new, unprotected data is created every day. This dynamic landscape makes it increasingly difficult for security teams to maintain control.

That's why modern data security must be adaptive. AI-driven solutions can continuously assess and 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 systems apply the right controls in real time. They analyze user behavior in context to reduce false positives and tailor policies based on shifting roles and access.

At the same time, AI-powered classification provides deeper visibility of sensitive data. It can automatically tag and label content to enforce encryption, access restrictions and editing rights. This adaptive approach enables security to scale with the business, without compromising protection.

Modern data security is comprehensive

A modern, comprehensive data security solution delivers unified visibility and control across all exfiltration channels, data sources and user types, whether the risk originates from carelessness, malicious intent or compromised accounts. A human-centric platform correlates user behavior with data movement across files, applications, cloud[13] environments and endpoints. As a result, it provides the context needed to distinguish between accidental actions, insider threats and external breaches.

This convergence of content, user and threat intelligence helps security teams to respond faster and more accurately. It ensures sensitive data is protected across the organization — from frontline employees to developers — within a single, integrated platform.

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

Beyond blocking risky activity, modern data security provides coaching to help them understand and change their behavior while keeping them productive.

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

Modern data security uses risk-aware controls to look for signs of compromise and apply additional security controls, blocking risky activity where needed.

Malicious users can intentionally exfiltrate data for personal gain.

Based on risk factors such as resignations or unusual activity around sensitive files, modern data security can monitor some users more closely, apply stronger access controls and proactively block malicious actions.

The case for a modern data security approach

Here's a closer look at how modern data security compares to legacy DLP in common use cases.

LEGACY DLP CAPABILITIES	MODERN DATA SECURITY	USE CASES
Monitor		
◐	●	Data classification
○	●	Insider risks (malicious users, privileged users, departing employees, server and workstation usage)
○	●	Third-party application usage
○	●	Threat hunting and DLP analytics (including data and file history)
◐	●	GenAI tool usage (application of acceptable use policies)
Detect		
●	●	Intellectual property and regulated data loss across cloud, email and endpoint (accidental data leakage or malicious data exfiltration)
○	●	Abnormal user behavior (compromised logins, using unapproved websites or malicious user behavior)
Prevent		
●	●	Data loss across channels (email, cloud and endpoint)
○	●	Abnormal user behavior (compromised or malicious users)
○	●	Misdirected emails and attachments
Respond		
○	●	Data loss, insider threat and account compromise investigations
◐	●	Integrations across SIEM, SOAR, business communication and ticket management tools

○ None ◐ Partial ● Complete

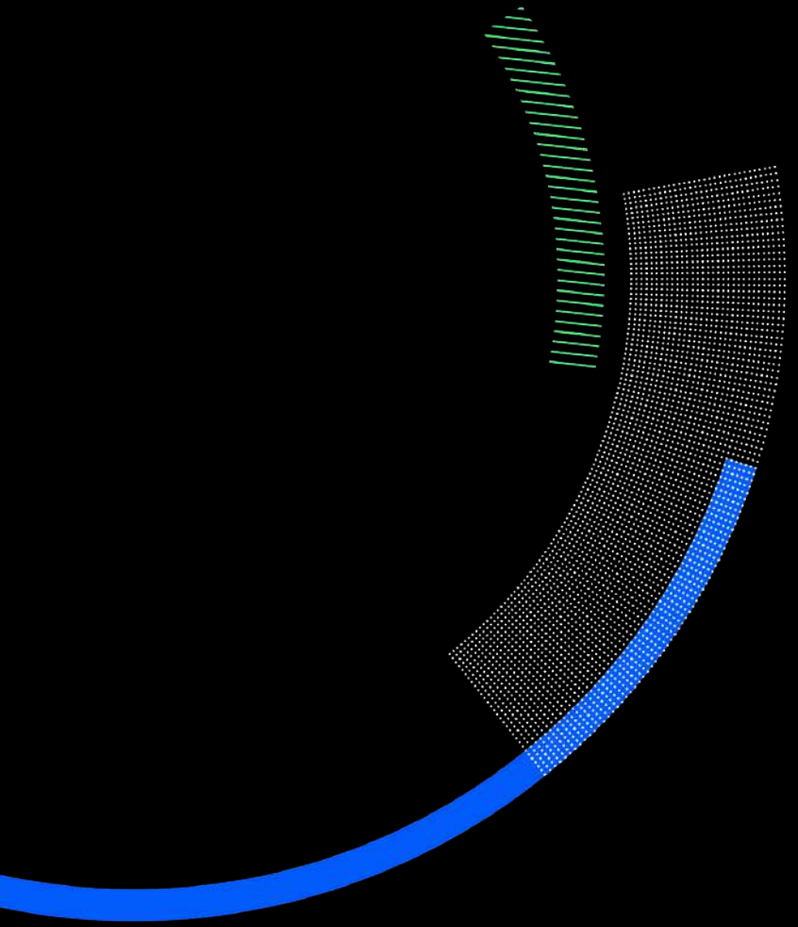
Conclusion

Today's organizations have embraced cloud-first strategies, hybrid work and rapid digital innovation. As a result, the risks to sensitive data have evolved — faster than legacy DLP solutions can keep up with. The reality is clear: traditional, content-centric tools are no longer sufficient in a world where data is dynamic, insider threats are rising and every user interaction carries context.

Modern 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 getting in the way of the business. 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 protection 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 →