Enterprise Strategy Group™ by TechTarget

Reinventing Data Loss Prevention: Adapting Data Security to the Generative AI Era

Todd Thiemann | Senior Analyst
ENTERPRISE STRATEGY GROUP

MARCH 2025

Research Objectives

Enterprises need to provide access to sensitive data while controlling against the unauthorized disclosure of that information from inadvertent leakage, insider threats, and outside attacks targeting data. Work-from-home and bring-your-own-device initiatives pose increased data loss prevention (DLP) challenges, and generative AI (GenAI) has opened new avenues for data leakage. Additionally, the proliferation of cloud services poses threats for data exfiltration, while intellectual property and trade secrets take new forms that do not lend themselves to conventional DLP solutions.

Although DLP is a top investment category when it comes to data security, enterprises continue to struggle to classify data and control against data loss. Whether an enterprise DLP solution or DLP functionality within another security technology, current offerings send considerable false positive alerts that distract teams that must evaluate and respond to alerts. Existing approaches focusing on regular expression (regex) rules are brittle and require considerable maintenance, while current DLP solutions frequently encounter scaling and performance issues. Furthermore, complex data types are difficult to categorize.

To gain insights into these trends, Informa TechTarget's Enterprise Strategy Group surveyed 370 IT and cybersecurity professionals at enterprise (i.e., 1,000 or more employees) organizations in North America (US and Canada) involved with identity security technologies and processes.

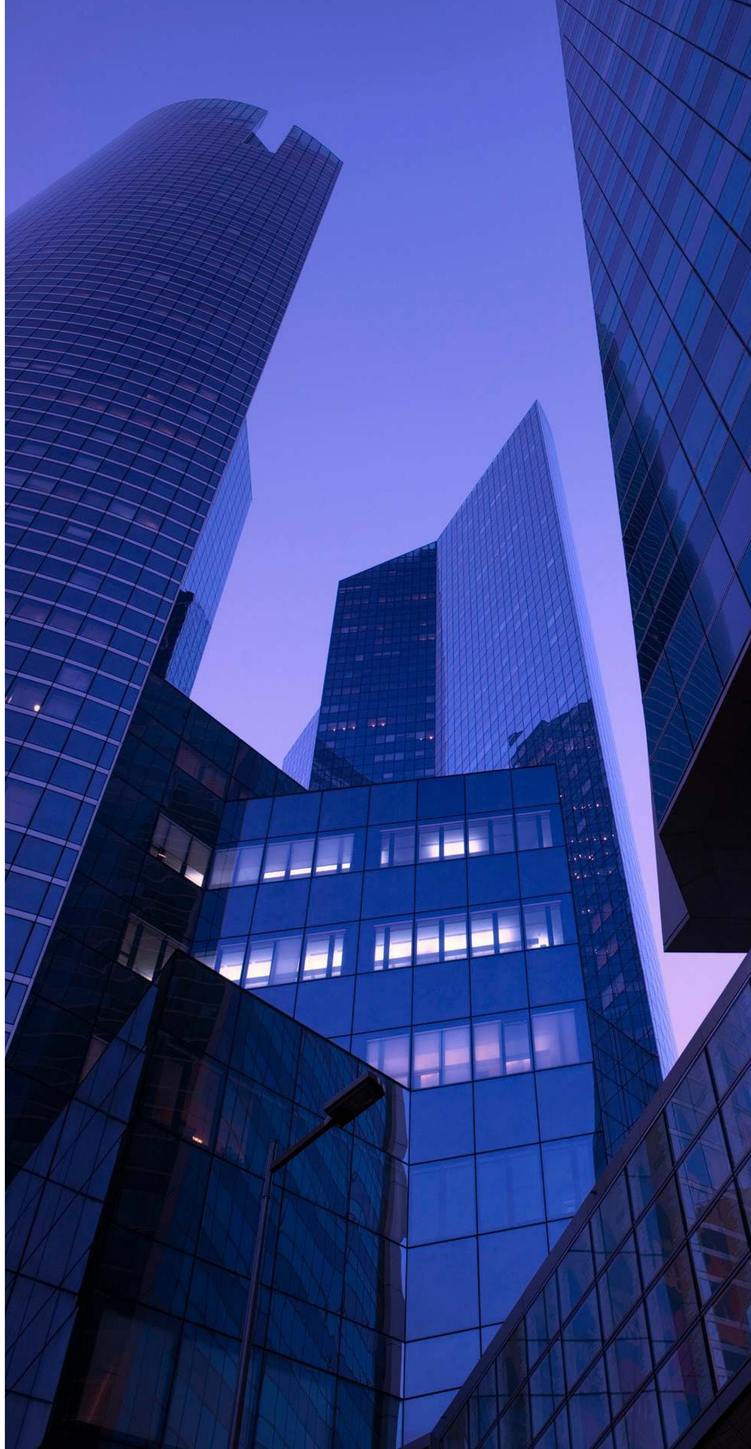
THIS STUDY SOUGHT TO:

Assess the state of the market for categorizing sensitive data and controlling against data loss across the enterprise attack surface.

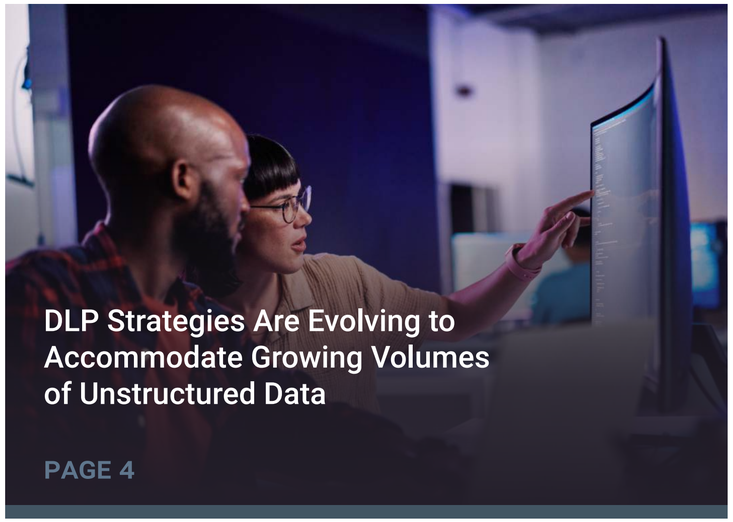
Explore the risk and management challenges posed by today's DLP solutions.

Uncover the challenges in controlling against unauthorized disclosure of sensitive data.

Highlight the emerging requirements for enterprises embracing new cloud services and GenAI technologies.



KEY FINDINGS

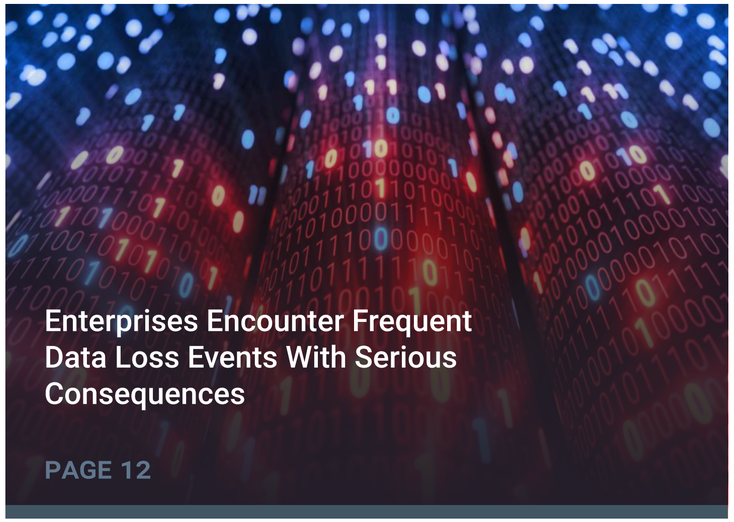


DLP Strategies Are Evolving to Accommodate Growing Volumes of Unstructured Data

PAGE 4

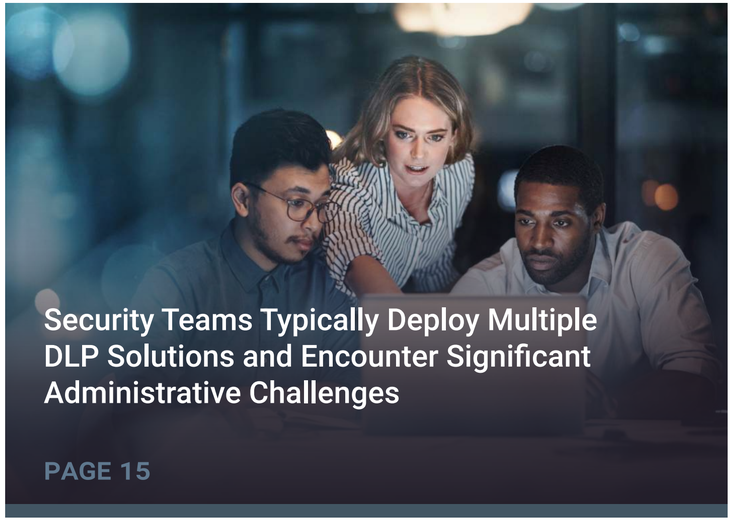
Data Loss Landscape Reveals Limited Visibility Into Large Volumes of Enterprise Data

PAGE 7



Enterprises Encounter Frequent Data Loss Events With Serious Consequences

PAGE 12



Security Teams Typically Deploy Multiple DLP Solutions and Encounter Significant Administrative Challenges

PAGE 15

Top DLP Priorities Include Reducing Alert Noise, Gaining Context Awareness, and Determining Risk Severity

PAGE 18

DLP Investments Are Growing and Changing to Streamline Workflows, Overcome Alert Noise, and Speed Remediation

PAGE 22

BACK TO CONTENTS

DLP Strategies Are Evolving to Accommodate Growing Volumes of Unstructured Data

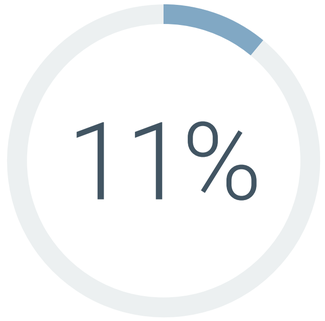
Most Enterprise DLP Strategies Are a Work in Progress

While DLP is a well-established solution area, enterprises continue to struggle with data loss. This may not be due to DLP technology, but rather how it is used. Indeed, only 38% of organizations believe they have a clear DLP strategy upon which they are executing. The balance of organizations are either refining and improving their existing DLP strategy or lack one entirely.

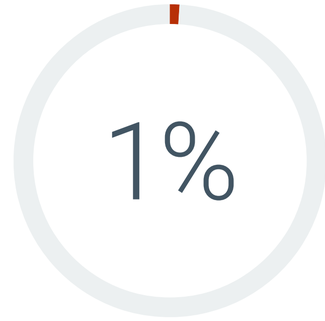
Status of DLP strategies.

Mature
We have a clear DLP strategy and are executing on it

Developing
We are refining and improving our DLP strategy



Inception
We are at the early stages of establishing our DLP strategy



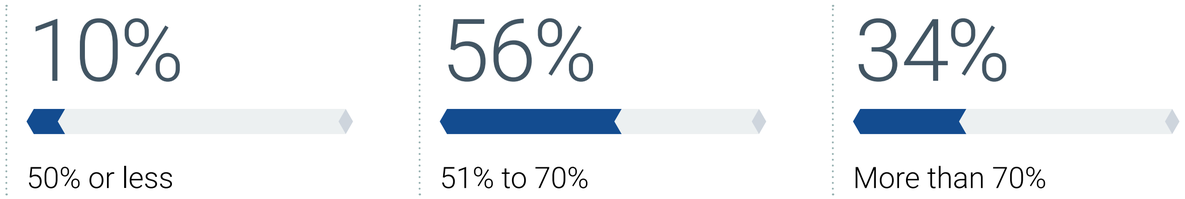
Ad hoc
We have no established DLP strategy

BACK TO CONTENTS

Unstructured Data Represents the Bulk of Enterprise Data, and Organizations Prioritize Accordingly

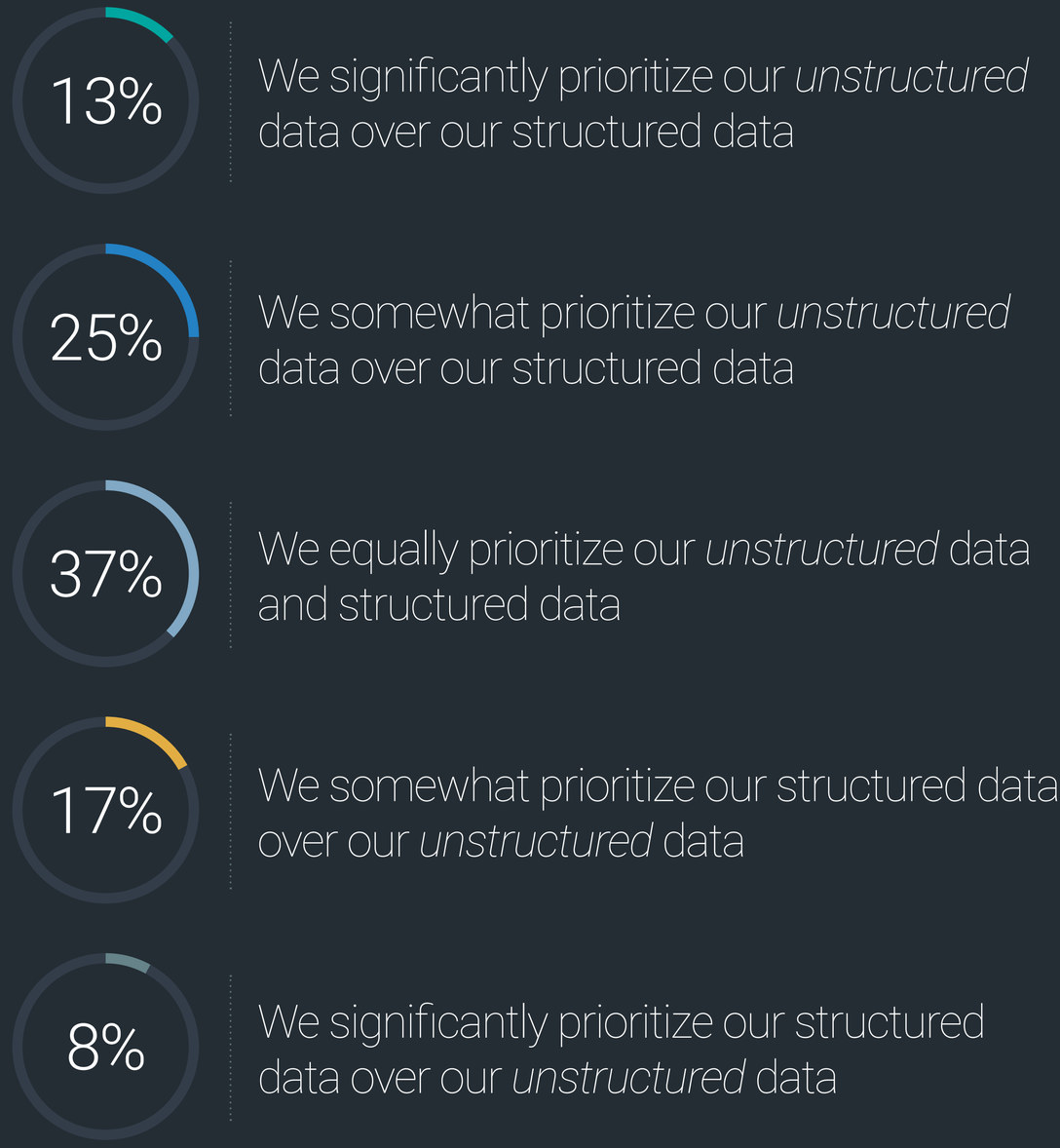
Unstructured data, which includes elements like email, spreadsheets, and text files, has no predefined format or organization, making it much more difficult to collect, process, and analyze. When asked about how much of their total data is unstructured, nine in ten organizations estimated the number to be more than half, which equates to about 64% of enterprise data on average.

Approximate percentage of total data that is unstructured data.

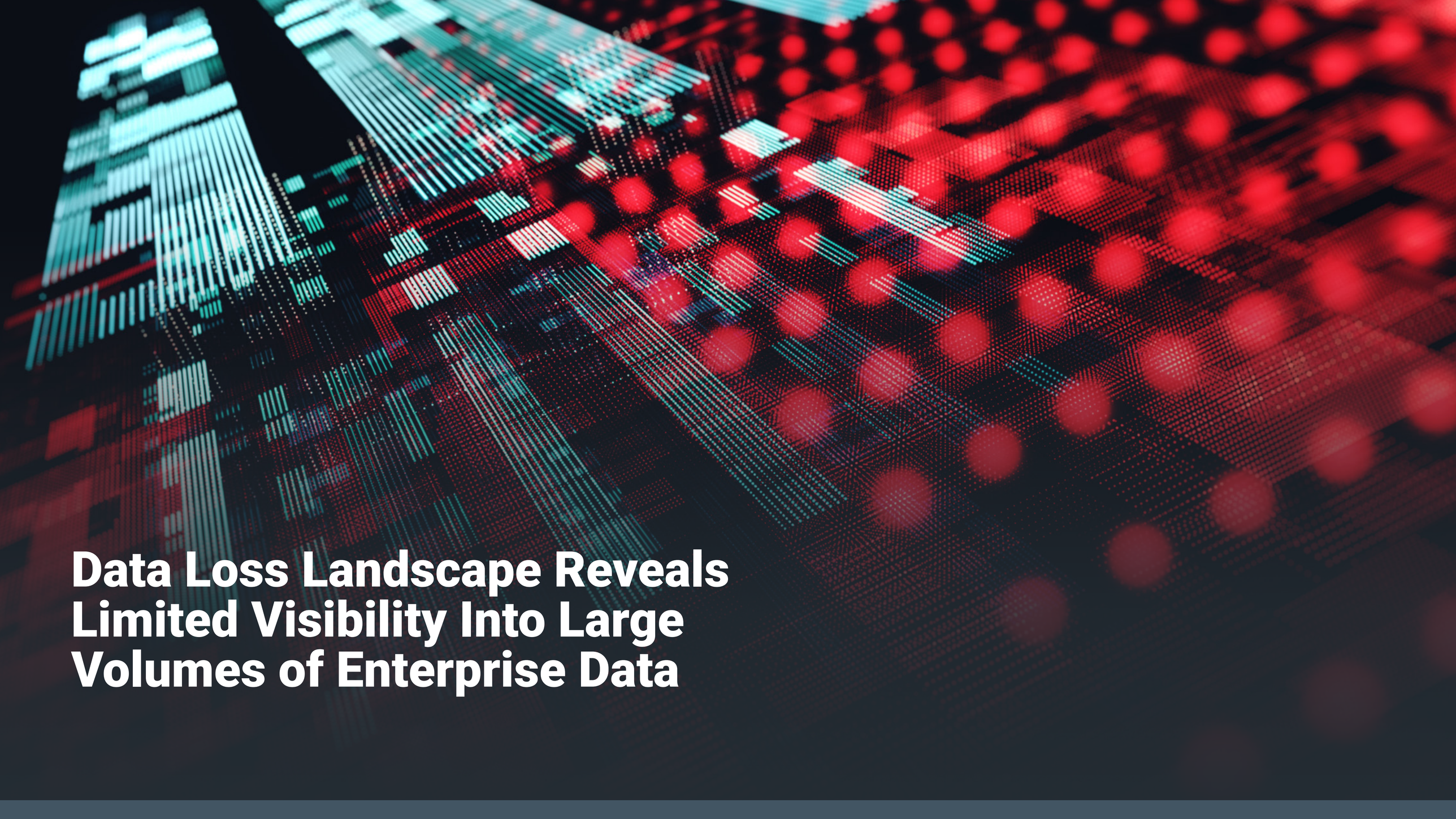


Given that sensitive data exists in both unstructured and structured data repositories, enterprises are concerned with securing both data types. While the plurality of organizations report equally prioritizing the two data types, more than one-third lean toward preventing unstructured data loss to some extent.

Types of data prioritized with DLP efforts.



BACK TO CONTENTS

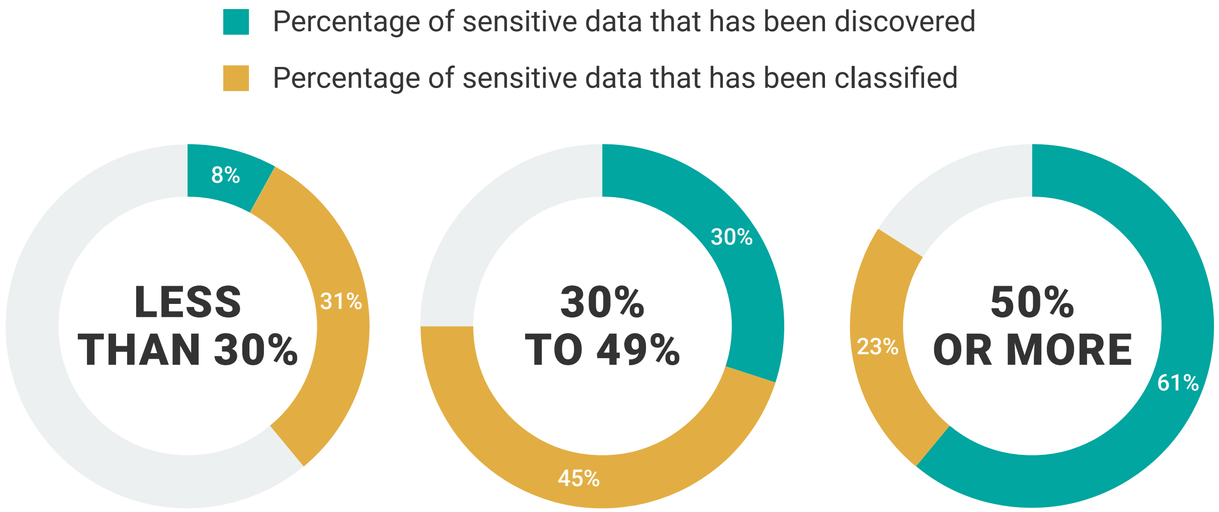


Data Loss Landscape Reveals Limited Visibility Into Large Volumes of Enterprise Data

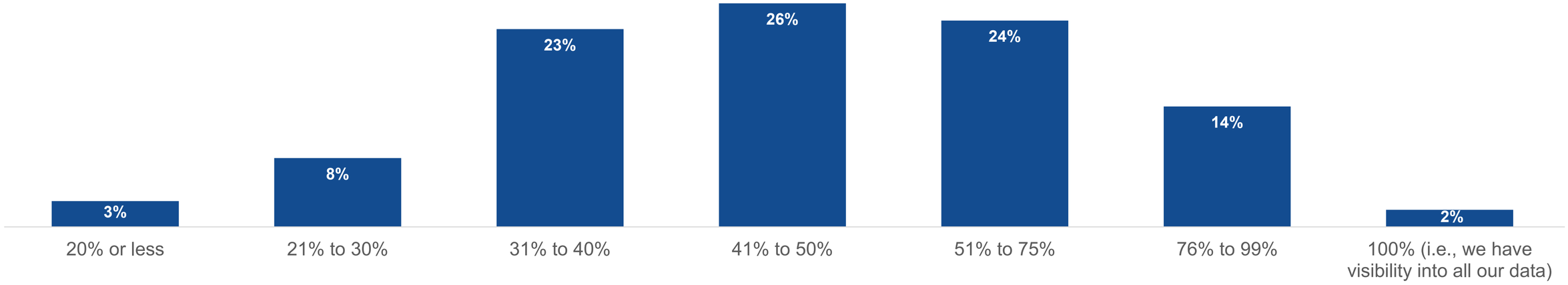
Enterprises Lack Visibility Into Large Swaths of Their Data Estate, Including Undiscovered and Unclassified Sensitive Data

Data discovery and classification are prerequisites to securing data and avoiding sensitive data loss. However, six in ten enterprises lack visibility into at least half of their data estate. The picture is even more bleak when it comes to sensitive data. An average of 56% of data was discovered, and 40% of the discovered data was classified. While the majority of organizations believe they've discovered at least half of their sensitive data, less than one-quarter have effectively classified the same amount. Clearly there is room for improvement when it comes to the visibility and organization of data.

Approximate percentage of sensitive data that has been discovered and classified.



Approximate percentage of total data organizations have visibility into.

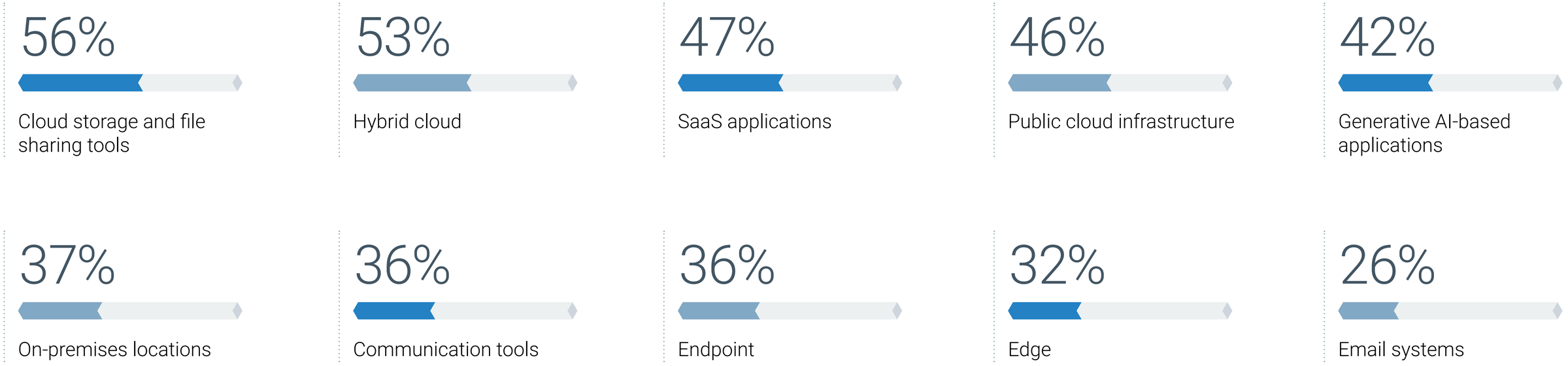


BACK TO CONTENTS

Sensitive Data Commonly Resides in Cloud Environments

Data repositories, infrastructure, and applications operating in cloud environments are most commonly identified as containing volumes of sensitive data. While endpoints and emails can provide the exfiltration vector, they are lower down the list of environments containing sensitive data. GenAI is the area to watch as enterprises deploy more large language model (LLM) infrastructure that uses sensitive data.

Environments in which sensitive data resides.

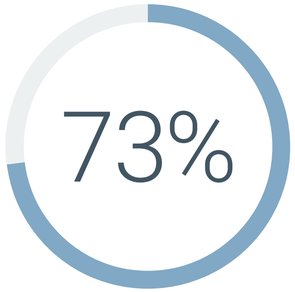


BACK TO CONTENTS

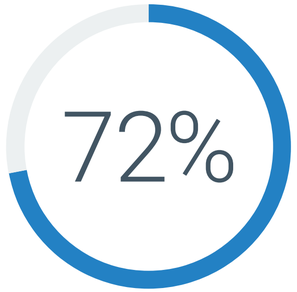
SaaS Clients and the Web Browser Are Key Control Points

Users access and share data using a variety of tools, each of which can be considered a control point. Nearly three-quarters of organizations indicate their users access and share sensitive data via SaaS clients such as Slack, Teams, Copilot, etc., (73%) and/or web browsers (72%). While organizations have tried to educate users about the types of information they disseminate through email, still six in ten report email clients as a medium for accessing and sharing sensitive data.

How users access and share sensitive data.

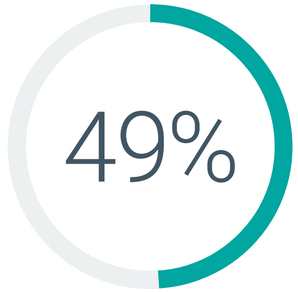


SaaS clients



Web browser

Email client



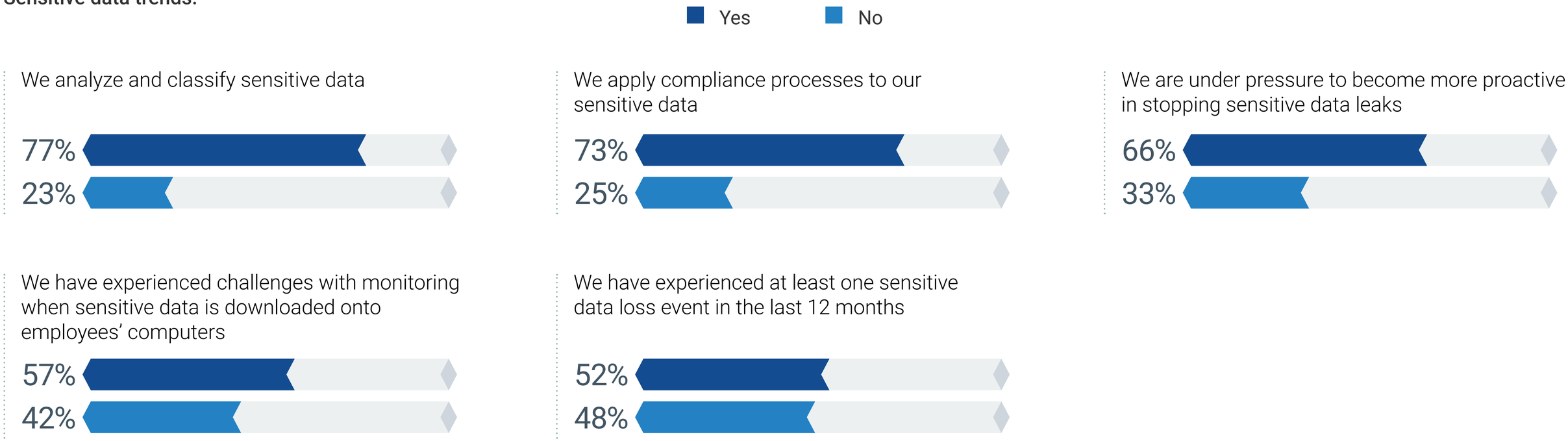
API

“Nearly three-quarters of organizations indicate their users access and share sensitive data via SaaS clients such as Slack, Teams, Copilot, etc., (73%) and/or web browsers (72%).”

Enterprises Are Under Pressure to Proactively Stop Leaks, but Half Have Experienced Recent Sensitive Data Leakage

While enterprises strive to do the basics, sensitive data leaks continue to occur. Although more than three-quarters (77%) of enterprises analyze and classify sensitive data, and another 73% apply compliance processes, 52% have experienced at least one sensitive data loss event in the past 12 months. It follows then that 66% report being under pressure to become more proactive in stopping sensitive data leaks.

Sensitive data trends.



BACK TO CONTENTS

Enterprises Encounter Frequent Sensitive Data Loss Events With Serious Consequences

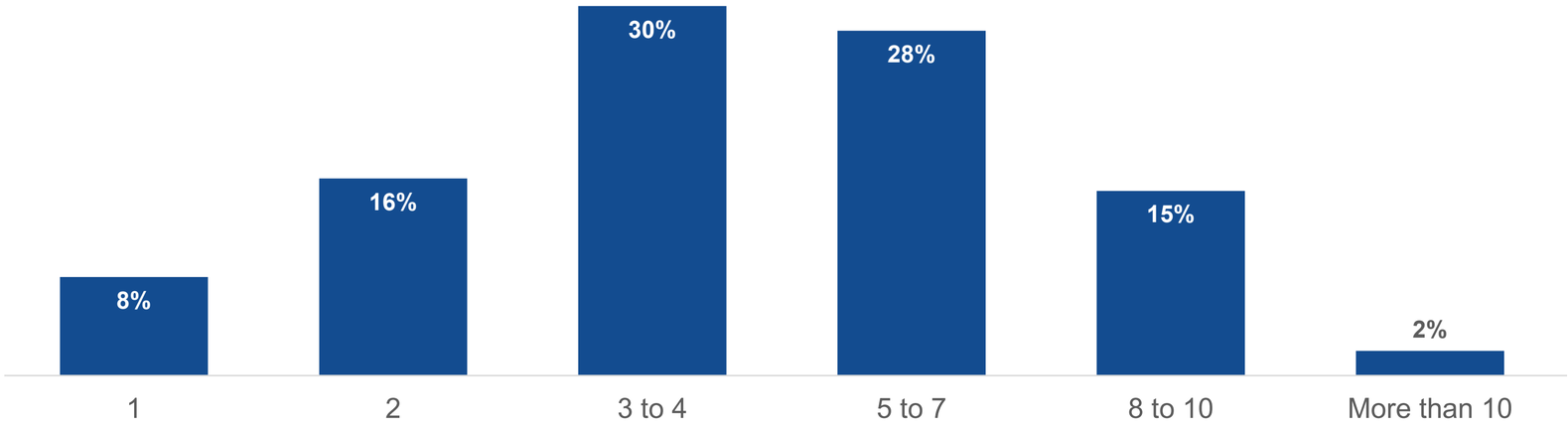
52%

have experienced at least one sensitive data loss event in the last 12 months.

Multiple Sensitive Data Loss Events Are the Norm

As seen previously, more than half of enterprises have experienced a sensitive data loss event in the last 12 months. Even more worrisome is that the vast majority of these victims indicated they suffered *multiple* data loss events, with nearly half (45%) reporting it happened at least five times.

Number of sensitive data loss events in the past 12 months.

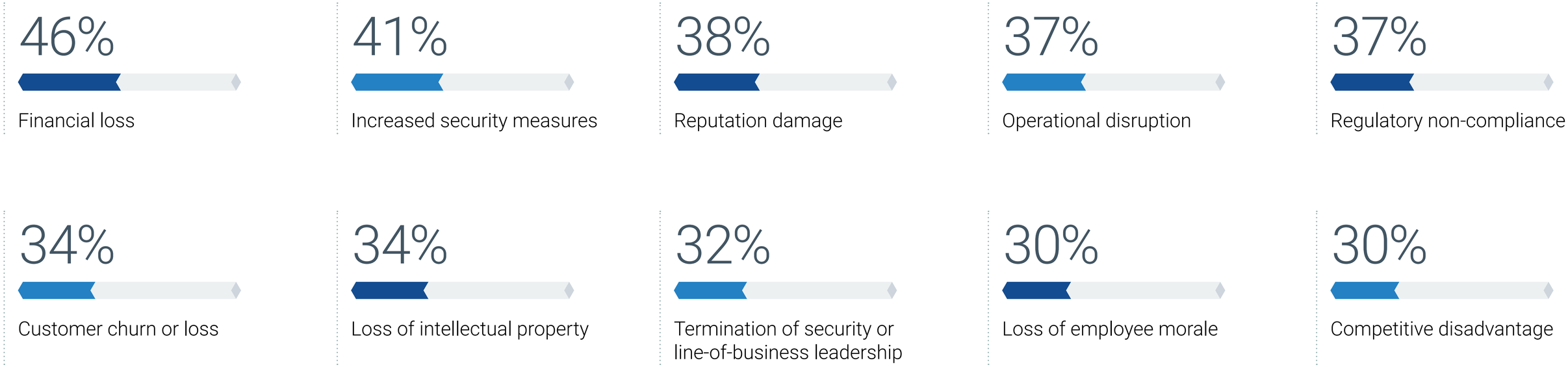


BACK TO CONTENTS

Organizations Suffer Manifold Business Impacts From Sensitive Data Loss

Nearly half of those organizations that experienced a sensitive data loss event within the past year attribute financial loss to the incident(s). Data loss causes damage in multiple other significant ways, including reputational damage, disrupted operations, lost customers, and lost intellectual property.

Business impacts of sensitive data loss events from the last 12 months.



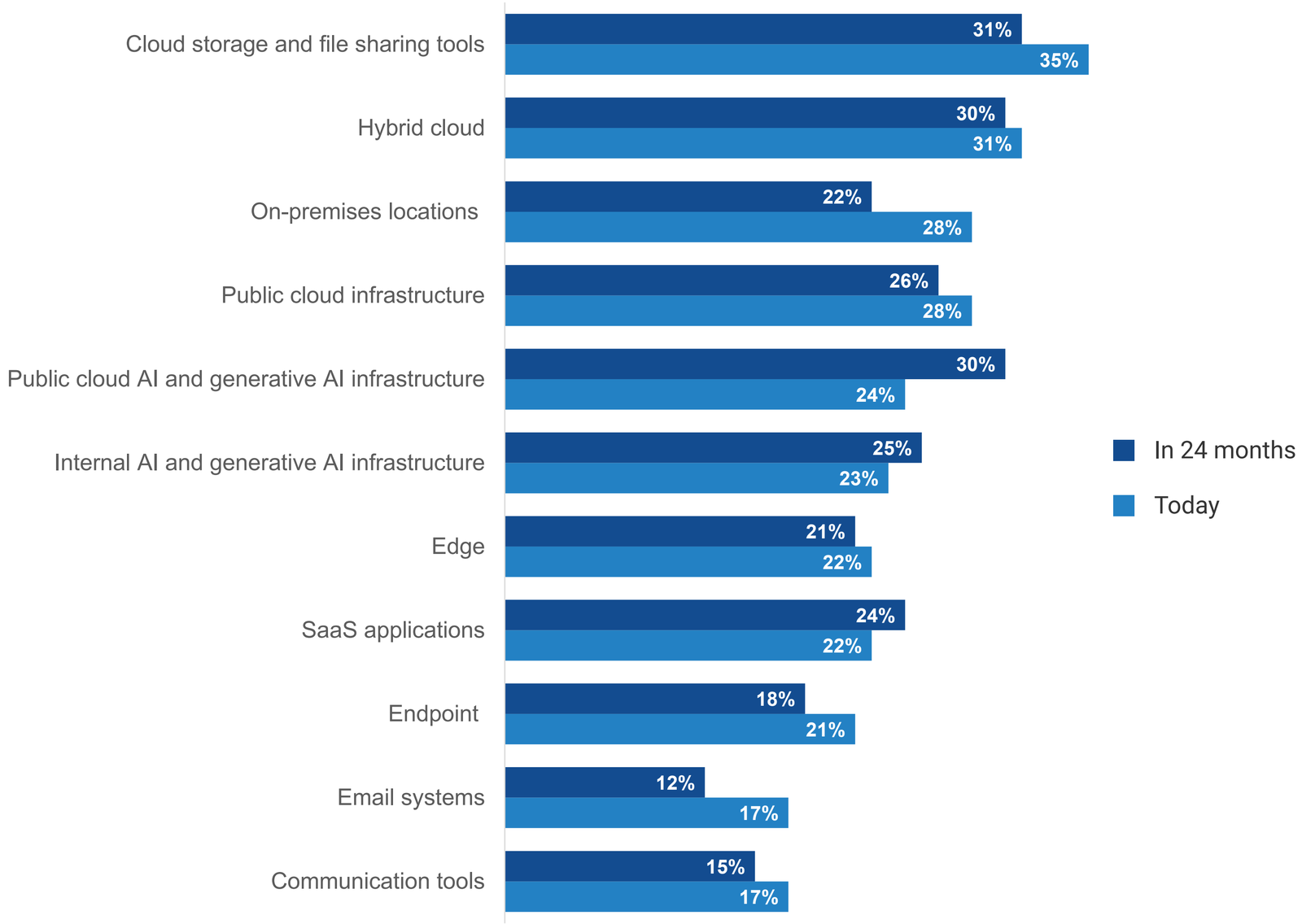
BACK TO CONTENTS

Security Teams Typically Deploy Multiple DLP Solutions and Encounter Significant Administrative Challenges

GenAI and SaaS Applications Are Expected to Grow in Importance for DLP

While cloud storage and hybrid cloud are most commonly deemed as important today for data loss prevention, GenAI, both internal and public cloud AI, along with SaaS application environments are expected to increase in DLP importance over the next 24 months.

Most important environments for data loss prevention.



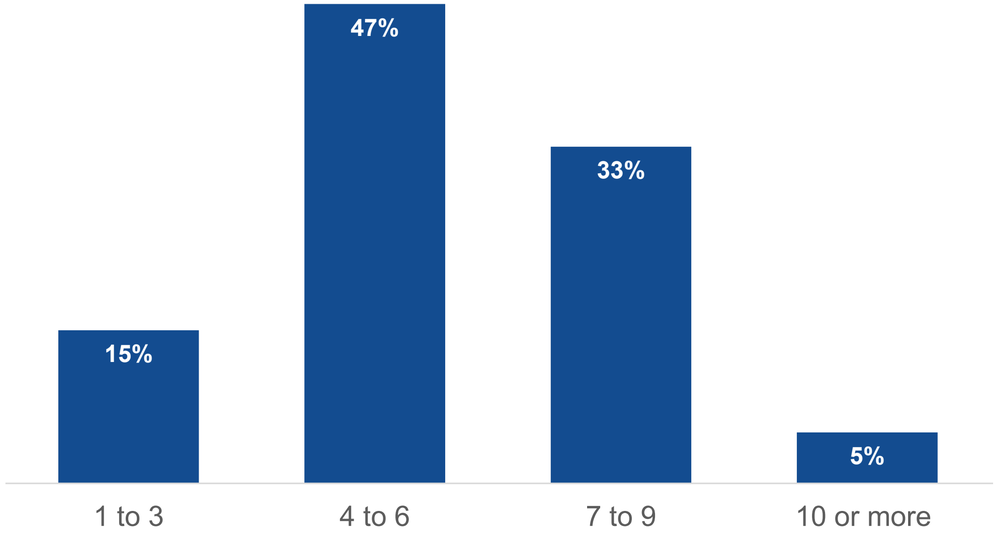
BACK TO CONTENTS

Enterprises Are Taking a Portfolio Approach to DLP, Causing Administration and Maintenance Challenges

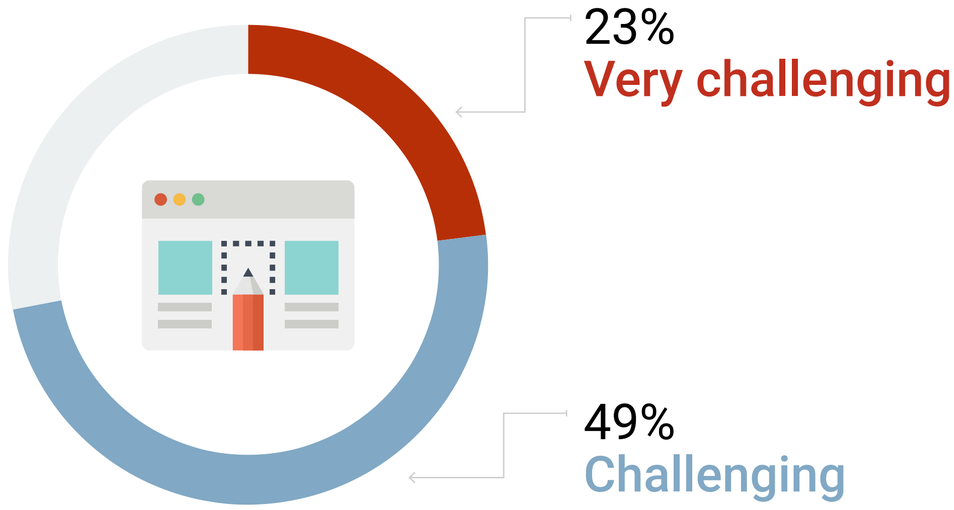
Many organizations use a combination of discrete DLP tools integrated at the endpoint, email, network, and cloud layers, as well as other tools with integrated DLP functionality across their entire environment. When asked to quantify their tool use, enterprises estimated using, on average, six DLP tools, with 85% leveraging at least four.

Given this DLP tool sprawl, it is not surprising that nearly three-quarters of enterprises find it challenging (49%) or very challenging (23%) to administer and maintain their existing DLP technology solutions and policies.

Number of data loss prevention tools in use.



Level of difficulty administering and maintaining existing DLP technology solutions and policies.



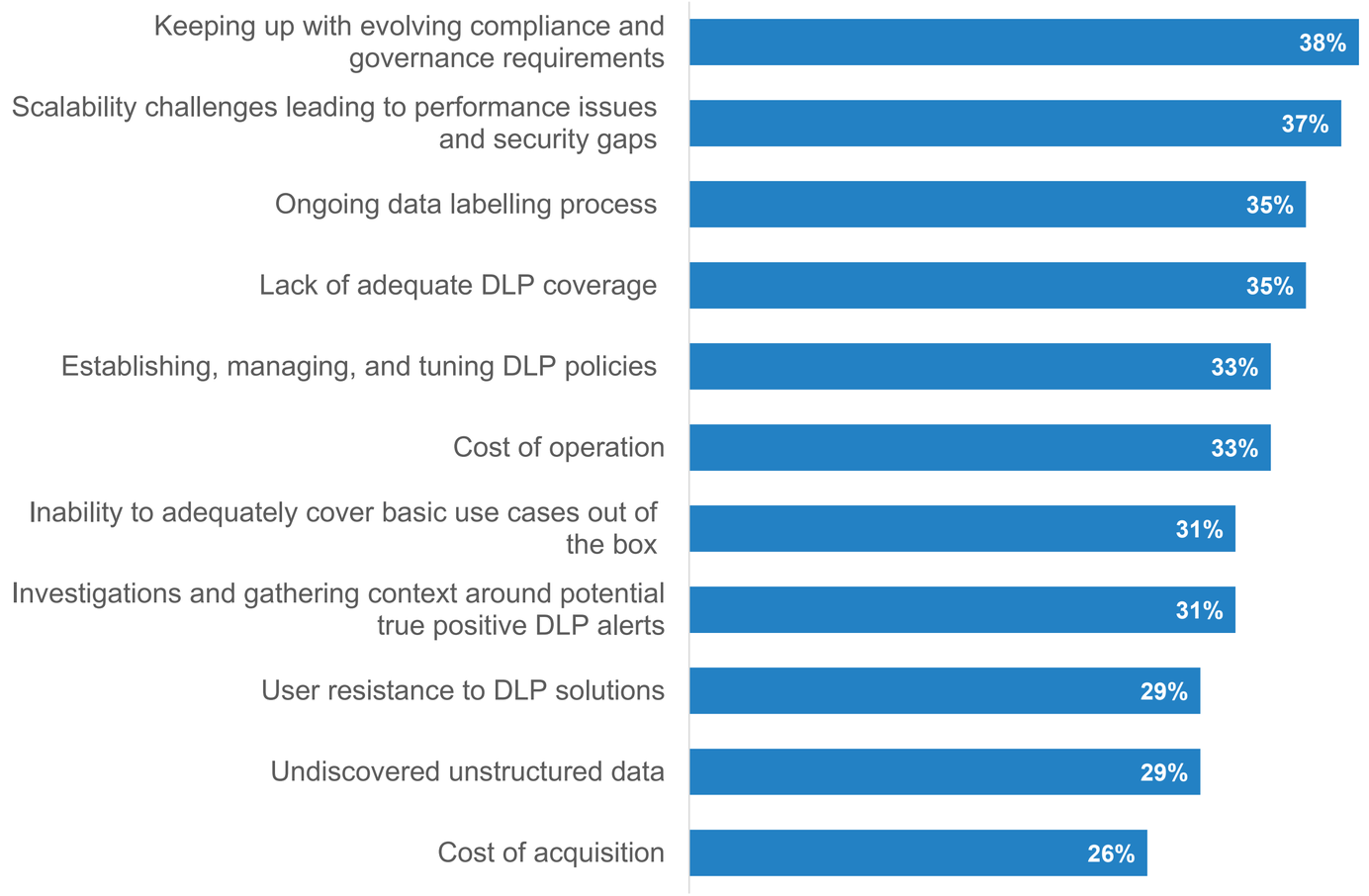
BACK TO CONTENTS

Top DLP Priorities Include Reducing Alert Noise, Gaining Context Awareness, and Determining Risk Severity

Organizations Face Many Points of DLP Frustration

What aggravates enterprises when it comes to DLP technology? The most common areas of frustration include maintaining compliance in a changing environment and scalability challenges leading to performance issues and security gaps. While cost of operation is on the minds of one-third of organizations, cost of acquisition is not as significant as a point of frustration.

Biggest frustrations with DLP technology.



BACK TO CONTENTS

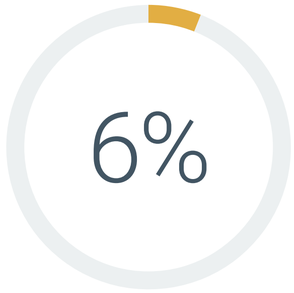
Alert Noise Pollution Is a Painful Problem

DLP solutions produce a large volume of alerts, many of which are false positives. Triaging and investigating DLP alerts drains staff time and resources. As such, 92% of enterprises think it is either important or very important to effectively reduce DLP alert noise.

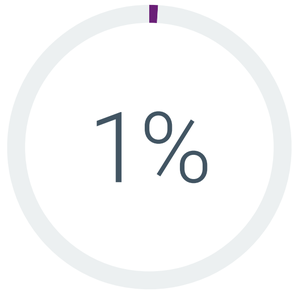
Importance of reducing alert noise produced from DLP controls.

Very important

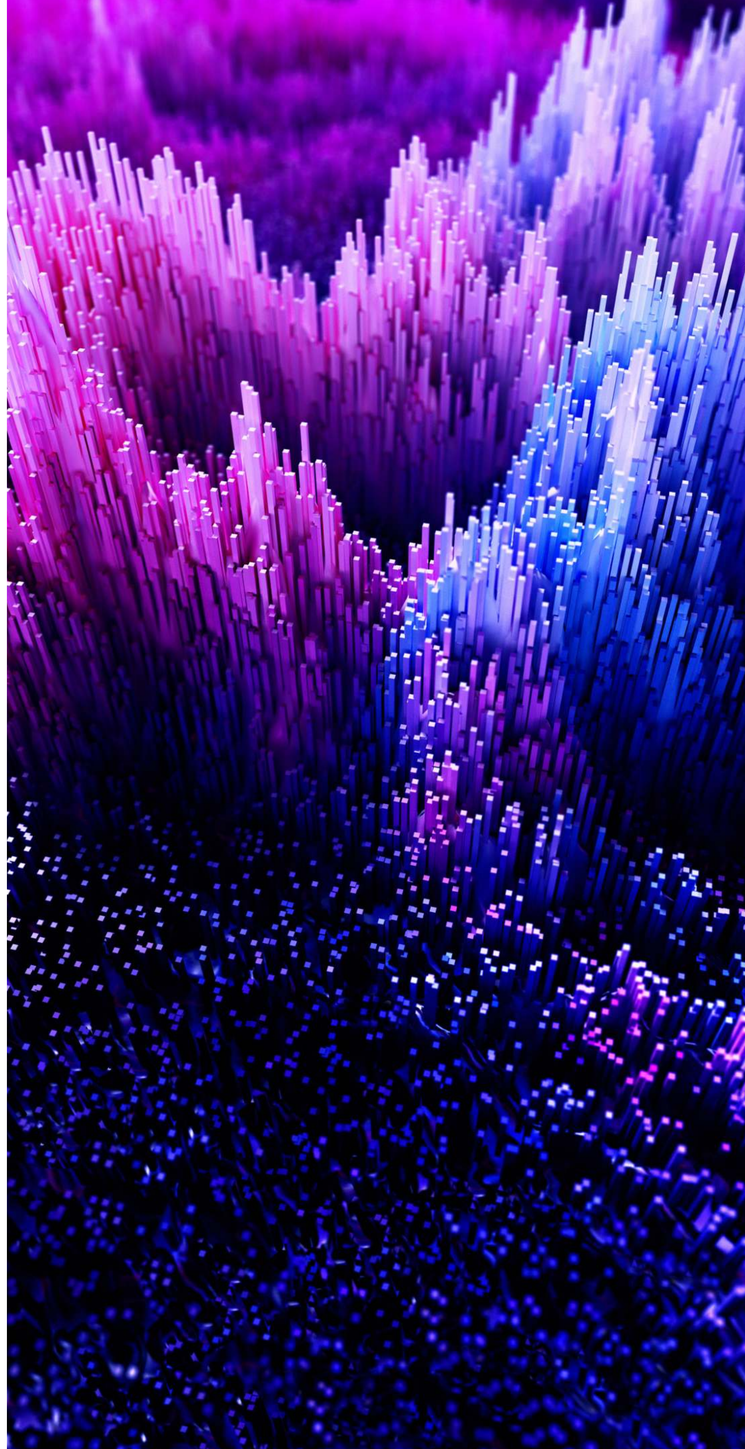
Important



Somewhat important



Not at all important



Funding Constraints and Compliance Preoccupations Lead Challenges With DLP Controls

While resource constraints for investigating and remediating alerts are the most commonly cited challenge experienced with the controls in place for preventing data loss, the biggest challenges lie in funding constraints to invest in better data leak controls along with a compliance focus that comes at the expense of data security best practices.

Challenges experienced with the controls in place for preventing data loss.

BACK TO CONTENTS

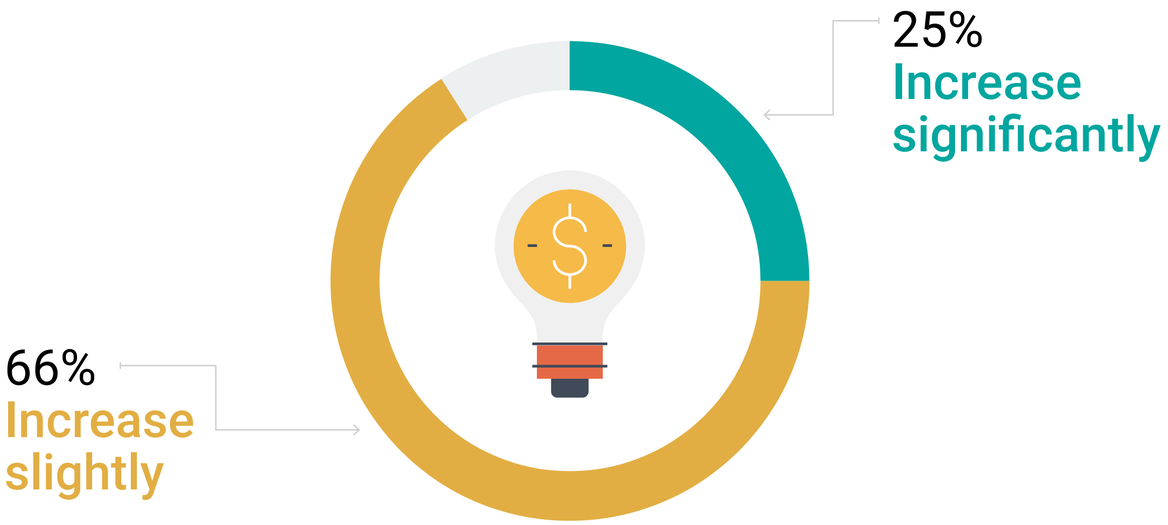
DLP Investments Are Growing and Changing to Streamline Workflows, Overcome Alert Noise, and Speed Remediation

DLP Budgets Are Primed to Grow Relative to Other Areas, Though Approaches to DLP Strategies Vary

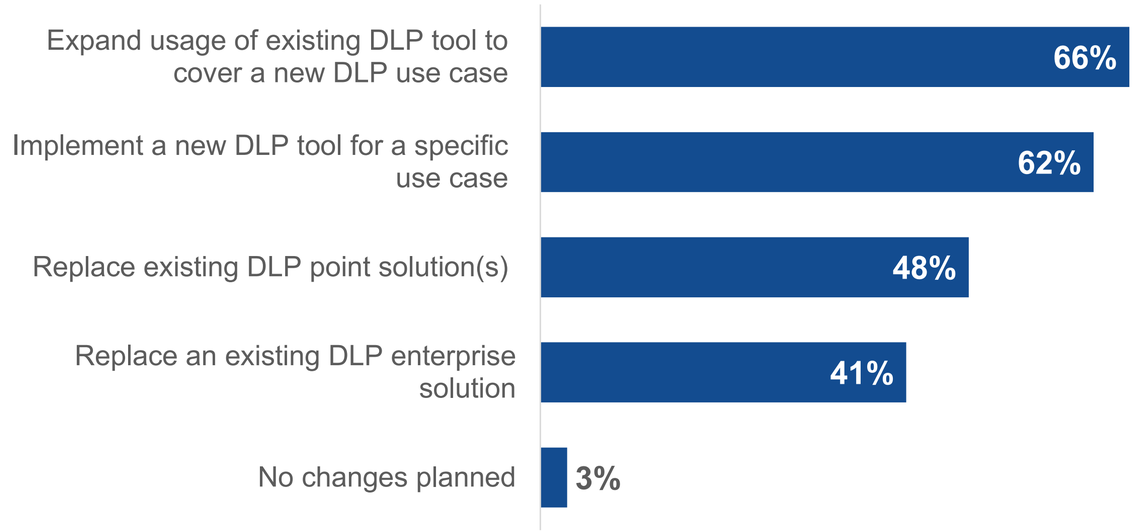
Relative to other areas of IT and cybersecurity, more than nine in 10 enterprises intend to increase spending on DLP solutions over the next 12 months, including 25% expecting this increase to be significant.

Expect notable change as the DLP space evolves in the next 12-18 months. Enterprises are striving to improve their DLP programs in divergent ways, with 66% expanding use of an existing tool, 62% deploying a new tool, and an astounding 40%+ either replacing existing point solutions or replacing an enterprise DLP solution.

Expected change in spending for DLP solutions over the next 12 months.



How organizations are looking to evolve DLP programs in the next 12-18 months.



BACK TO CONTENTS

harmonic

ABOUT

Harmonic Security gives security teams the tools to prevent sensitive data going into GenAI without the headaches of labeling and complex rules. Their pre-trained data protection models enable secure innovation through user interaction and gentle nudges.

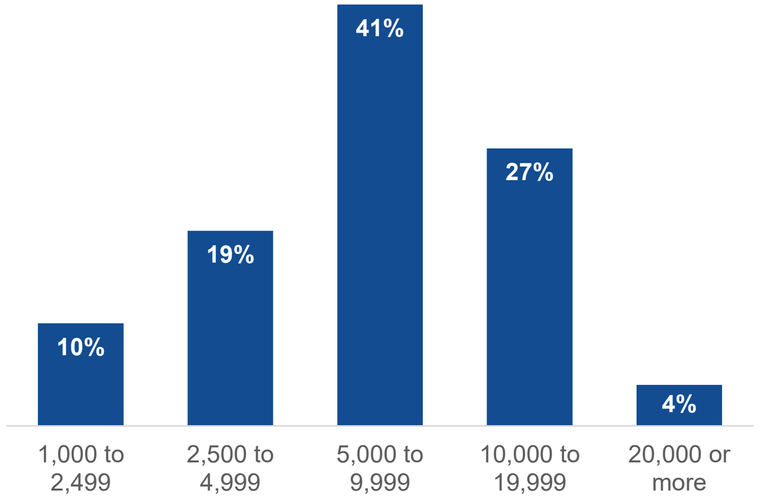
LEARN MORE

RESEARCH METHODOLOGY AND DEMOGRAPHICS

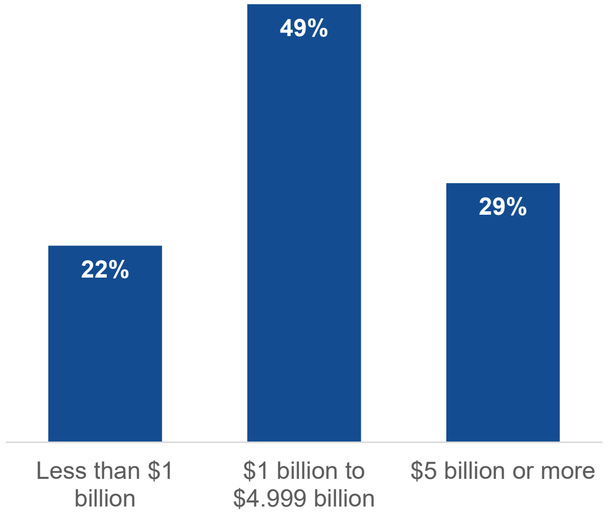
To gather data for this report, Enterprise Strategy Group conducted a comprehensive online survey of IT and cybersecurity professionals from private- and public-sector organizations in North America between December 19, 2024, and January 7, 2025. To qualify for this survey, respondents were required to be involved with identity security technologies and processes. All respondents were provided an incentive to complete the survey in the form of cash awards and/or cash equivalents.

After filtering out unqualified respondents, removing duplicate responses, and screening the remaining completed responses (on a number of criteria) for data integrity, we were left with a final total sample of 370 IT and cybersecurity professionals.

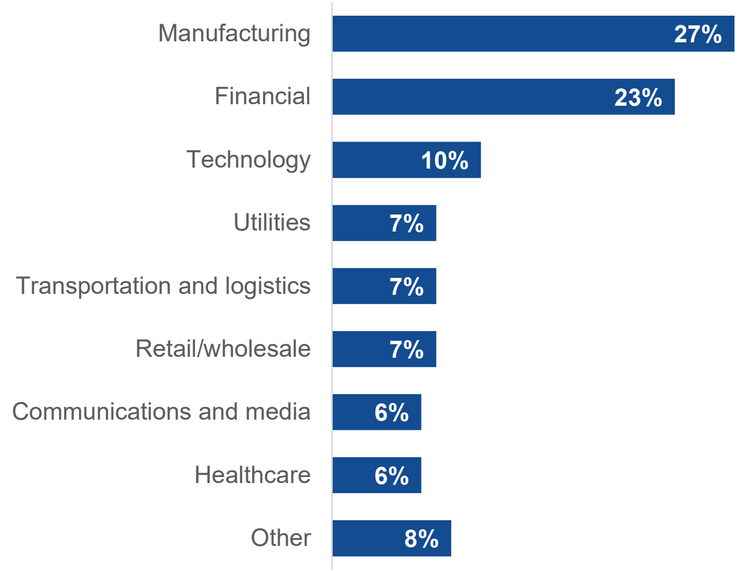
Respondents' organizations by number of employees.



Respondents' organizations by annual revenue.



Respondents' organizations by industry.



BACK TO CONTENTS

Enterprise Strategy Group is an integrated technology analysis, research, and strategy firm providing market intelligence, actionable insight, and go-to-market content services to the global technology community.