Vanguard
Report

September 2023

Securing More Data in More Places With Sensitive Data Discovery and Classification in the Cloud

Commissioned by

CYERA

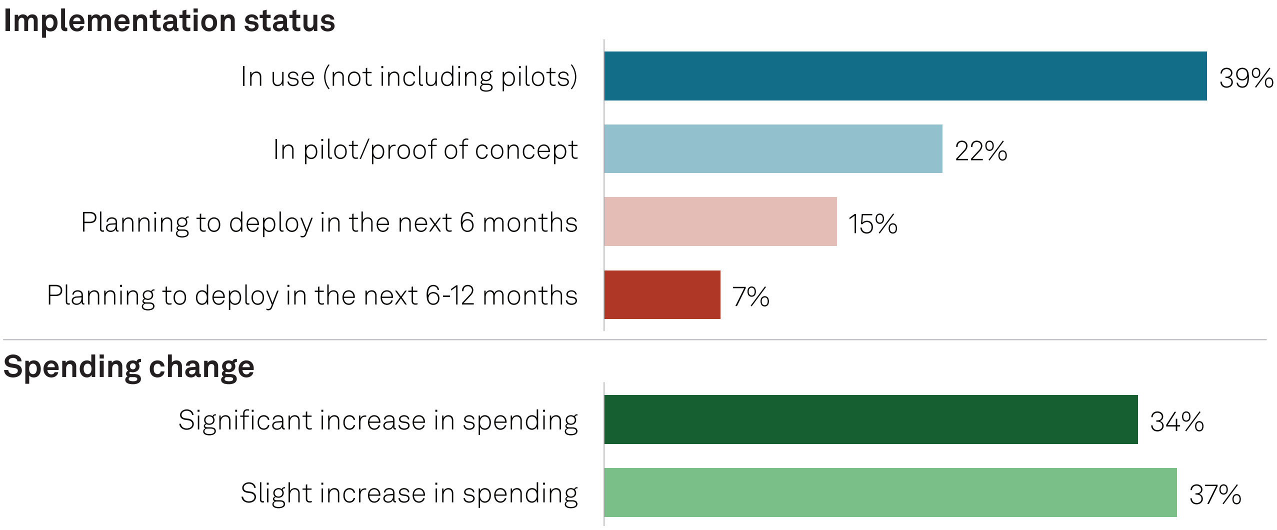
S&P Global
Market Intelligence

Introduction

Sensitive data discovery and classification is the top-cited security technology that managers plan to implement in the next 12 months, according to 451 Research's Voice of the Enterprise: Information Security, Technology Roadmap 2023. At the same time, our research shows that the speed and proliferation of cloud-native development is greater than ever. Use of cloud-native tools is widespread, with a strong pipeline for adoption. More than half of organizations surveyed (52%) indicate that they have cloud-native technologies and methodologies in use, according to our Voice of the Enterprise: Cloud Native, Adoption and Usage 2023 study. Another 14% say they are in discovery or proof of concept, and 14% are planning to implement within the next 12 months.

According to the Information Security survey cited above, only 39% of respondents are using sensitive data discovery and classification, although 22% are in pilot/proof of concept and another 22% plan to deploy it in the next 12 months, indicating significant growth in the segment. Among those who have the technology in use, 71% say they will increase their spending on it in the next 12 months.

Figure 1: Implementation status and spending change — sensitive data discovery and classification



Q. What is your organization's status of implementation for the following information security technologies? - Sensitive data discovery and classification.
Base: Respondents whose organizations have at least one information security technology in use (n=148).
Q. How will your spending on sensitive data discovery and classification change in the next 12 months?
Base: Respondents whose organization has sensitive data discovery and classification technology in use (n=41).
Source: 451 Research's Voice of the Enterprise: Information Security, Technology Roadmap 2023.

A major roadblock to sensitive data discovery and classification is implementation, which typically involves a highly manual process that organizations may never achieve. This presents a significant issue, given the rapidly increasing pace of cloud deployments. In this report, we look further into challenges of data security, vendor developments that can aid implementation and various strategies.

The Take

The nature of cloud-native datasets leads to rapid creation followed by persistence using inexpensive storage. For example, DevOps methodologies have increased the quantity of deployments through automation of continuous integration/continuous delivery processes, and newly generated data is collected and rapidly shared through the software supply chain. If the process is not managed properly, data security risks increase with every new application release. Organizations must clearly understand the risks their data presents and the processes required to understand and reduce those risks. They must support their stakeholders' objectives to grow their business while proactively managing the risk of data breaches. Gaining a firm understanding of where sensitive data resides and how it is used needs to be a major component of this strategy.

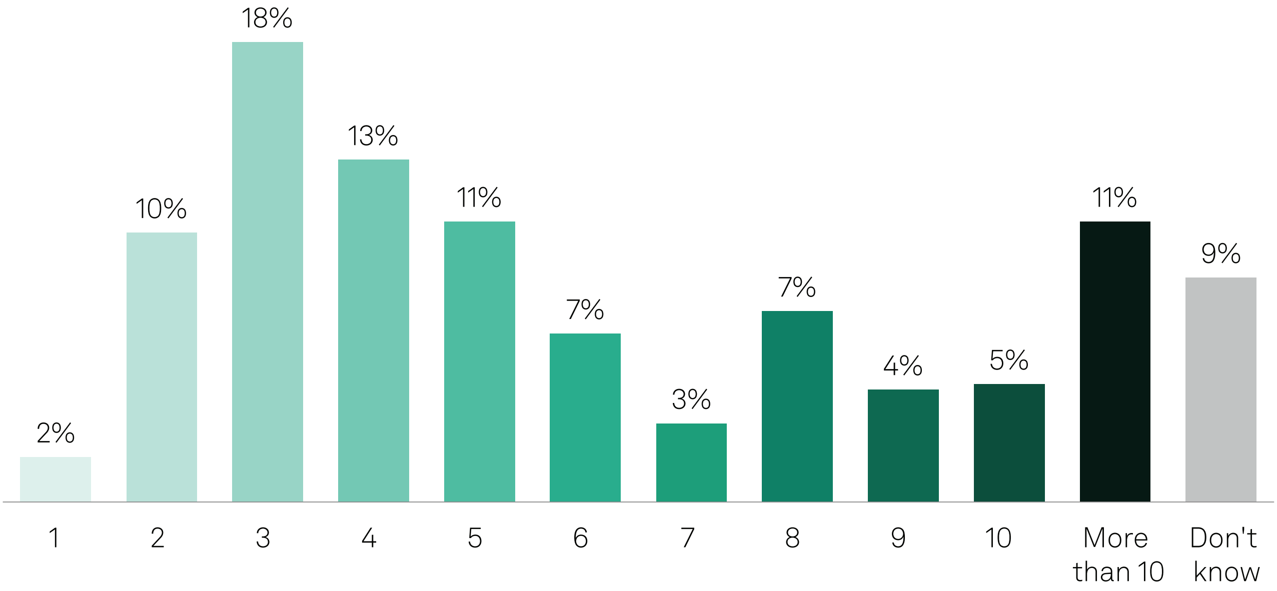
To better prioritize risk mitigation, enterprises need to understand their current data security risk posture and also be prepared to guard against risks as they emerge. Fully integrating security policy and compliance mandates within core DevOps methodologies enables risk management, compliance, and security operations personnel to offer more proactive guidance to line-of-business developers. However, managing data content alone does not provide a complete picture of risk. A fully realized data security risk management strategy must consider the conditions, processes and identities that work with and have access to sensitive data. In turn, policy creation and adoption must also continuously refine how sensitive data is defined, interpreted and implemented.

Use cases

The landscape

More data, more places. According to the Voice of the Enterprise: Storage, Data Migration 2023 study, enterprises have on average 5.21 copies of their business data. Much of that data is stored in the cloud, with 85% of respondents indicating they have cloud storage in use, in pilot/ proof of concept or in plan for the next 12 months. And the majority of data is in motion: Two-thirds (66%) of respondents indicate that on-premises data is transferred to or from public cloud environments on a continuous, daily or weekly basis. Keeping track of all that data is clearly a daunting task.

Figure 2: Copies of business data across the organization



Q. How many copies of your business data exist across your organization, both on-premises and in the cloud (e.g., primary storage, backups, archives)?
Base: IT decision-makers whose organizations adopt or plan to use cloud storage (n=458).
Source: 451 Research's Voice of the Enterprise: Storage, Data Migration 2023.

More cloud-native applications. There are many cloud-native applications, with more on the way. When asked what portion of their applications are architected using cloud-native technologies, 59% of respondents say more than 50% of applications are cloud-native. Two years from now, 77% of respondents expect more than 50% of applications will be cloud-native.

Privacy and compliance difficulties. Privacy and compliance remain challenging. In 451 Research's Voice of the Enterprise: Data & Analytics, Data Governance and Privacy 2022 study, respondents indicate that the complexity or volume of data sources is their top challenge to managing data privacy. In the same study, respondents cite improved regulatory and legal compliance as their No. 2 objective in improving data governance and data privacy efforts.

It's not just IaaS and PaaS: SaaS consumption is also increasing. Organizations concerned with managing data security need to consider the increasing use of SaaS, which grew during the pandemic due to increased remote work. Mobility/remote work enablement is the top-cited benefit of SaaS implementation in 451 Research's Voice of the Customer: Macroeconomic Outlook, SME Tech Trends, Cloud Adoption 2022 study. Unsurprisingly, security/compliance is the second-biggest SaaS-related concern. SaaS environments are difficult to manage from a data security perspective since SaaS consumers must trust the service provider's data security controls, and independent monitoring can be difficult.

We're all agile. Agile and DevOps methods are becoming the de facto work pattern. According to Voice of the Enterprise: DevOps, Developer Experience 2023, 84% of enterprises have some level of DevOps implementation, with 31% adopting DevOps exclusively, and we expect these percentages to grow. Half (51%) of respondents deploy software applications to production weekly or more often, amplifying the need for data security platforms that can keep pace. Today's organizations are more cloud native and have more data, and agile and DevOps methods add to the complexity, putting manual data security efforts at odds with the pace of business requirements.

The process

Security requires a livestream, not just a selfie. Security and compliance initiatives can no longer be conducted as point-in-time exercises. While regulations such as PCI DSS were the first to incorporate technical controls such as data encryption, activity monitoring and segmented networks, many implementations are based on the minimum regulatory requirement, which is typically a periodic report or assessment. While point-in-time checks are better than nothing, compliant does not mean secure.

Lean into automation. With the speed and complexity of today's data security landscape, automation is required to keep pace. For enterprises to understand and manage the multiple, constantly moving copies of data they possess in the context of production code that changes every week and operations that span multiple privacy and regulatory jurisdictions, automation is essential — not only to discover existing data security risks but to continually discover new risks as they arise.

Lean into autonomy. Underlying much of the complexity is the sheer variety of applications, databases and other repositories whose operation cannot be interrupted for security assessments. In many cases, these applications, platforms and other third-party offerings are changing as rapidly as enterprises' own code. Data security tools must consider how to operate and integrate in these dynamic, continuously available environments while creating minimal friction to production systems. This means discovery and classification processes must be autonomous and agentless to account for unknowns, such as data stores that the security team is unaware of, and out-of-band to eliminate the need to schedule performance-impacting scans and to ensure continuously evolving insights.

Consider stakeholders. Various stakeholders are increasingly motivated to improve data security. Consider “martech” operators that manage customer data platforms. According to the Voice of the Enterprise: Customer Experience & Commerce, Merchant Study 2022, 45% of merchants and marketers identify data security concerns as the No. 1 growth inhibitor for their business, above competition from retailers and marketplaces (35%) and logistics and supply chain issues (33%). However, in the same study, when respondents were asked which technologies their organization plans to increase investment in, data security lands in a three-way tie for fourth place, trailing categories such as CRM and customer support. While marketing and commerce teams may be familiar with marketing technology tools, they may be less familiar with security and governance tools.

An iterative approach is needed. Like many complex projects, sensitive data discovery and classification does not lend itself to a “boil the ocean” approach. A successful implementation might start with continuous (livestream) processing, as well as automation to enable autonomy by addressing accurate tagging of assets to enable both individuals and existing tools to function more effectively. This could be followed by policy/governance guidelines that inform and educate stakeholders to build more awareness into the security culture. Additionally, it is important to make remediation defensible.

Remediation must be defensible. Not every issue will allow for a simple fix. Understanding data security risks requires consideration of both content and context, as does risk remediation. While publicly exposed databases full of sensitive information present immediate problems, rash remediation such as blocking all database access may significantly harm business operations. Defensible remediation allows developers and operators to understand and incorporate better controls sooner while also incorporating other cloud security best practices such as secrets management, dynamic roles and ephemeral access tokens.

Conclusion

More diffusion. With an increasing number of third-party and N-party relationships, diverse stakeholders must better align to understand how their sensitive data is handled in relation to business goals. Increases in innovation must be accompanied by equal prioritization of security by stakeholders — in a proper business context. Data security concerns will only grow more diffuse as data's sensitivity and importance permeates all parts of the organization.

Embrace automation. Prioritize automation and data security time to value by investing in tools that continuously discover and automatically classify data, leveraging intellectual property (IP) along with AI and machine learning (ML) to understand your unique data and automatically notify, recommend and remediate issues across all environments, including on-premises and cloud. Layering advanced AI/ML with existing IP and custom logic (such as regular expressions and object definitions) is necessary to make discovery, classification and data risk assessments defensible as tools that security teams can use to keep pace with the way businesses create, consume and leverage data.

Better developer/operator experience. As a corollary to the diffuse nature of data security, developer and operator experience must be considered. The cloud-native and agile workflows they embrace must be complemented by cloud-native and agile controls. Done properly, sensitive data discovery and classification provides a variety of advantages including understanding and managing risks from data breaches, prioritization of remediation efforts, decreased compliance issues and improved productivity.

About the author

Justin Lam

Research Analyst

Justin Lam is a research analyst in the Information Security Channel at S&P Global Market Intelligence, focusing on data security. He has had a variety of roles within established and emerging data security vendors over the last 15 years, both in technical and commercial functions. Justin holds a BS in Business from Carnegie Mellon University and is based in the San Francisco Bay Area.

Mark Ehr

Senior Consulting Analyst

Mark Ehr is a senior consulting analyst in the S&P Global Market Intelligence Technology, Media & Telecommunications (TMT) Research team based in Denver. Prior to joining S&P Global, he spent 12 years at IBM in worldwide security sales enablement and QRadar SIEM product management roles.

Prior to IBM, he worked for BigFix, Cabletron, Enterprise Management Associates, Ping Identity, Polarsoft, Siebel Systems and Sybase, in a variety of roles including consultant, entrepreneur, industry analyst, product marketer, software developer and tech seller.

Mark holds a bachelor's degree in Computer Science from Metropolitan State University of Denver.

About S&P Global Market Intelligence

At S&P Global Market Intelligence, we understand the importance of accurate, deep and insightful information. Our team of experts delivers unrivaled insights and leading data and technology solutions, partnering with customers to expand their perspective, operate with confidence, and make decisions with conviction.

S&P Global Market Intelligence is a division of S&P Global (NYSE: SPGI). S&P Global is the world's foremost provider of credit ratings, benchmarks, analytics and workflow solutions in the global capital, commodity and automotive markets. With every one of our offerings, we help many of the world's leading organizations navigate the economic landscape so they can plan for tomorrow, today. For more information, visit www.spglobal.com/marketintelligence.

CONTACTS

Americas: +1 800 447 2273
Japan: +81 3 6262 1887
Asia Pacific: +60 4 291 3600
Europe, Middle East, Africa: +44 (0) 134 432 8300

www.spglobal.com/marketintelligence
www.spglobal.com/en/enterprise/about/contact-us.html

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global keeps certain activities of its divisions separate from each other to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain nonpublic information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its websites, www.standardandpoors.com (free of charge) and www.ratingsdirect.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at www.standardandpoors.com/usratingsfees.