LEARNING MADE EASY

Laminar Special Edition

Data Security Posture Management (DSPM)

for dummies®
A Wiley Brand

Keep your data secure

Eliminate shadow data

Get visibility and control

Brought to you by

Laminar

Lawrence Miller

About Laminar

Laminar is the leading agile data security platform and provides organizations the visibility and control they need to achieve data security, governance, and privacy in the cloud. Our cloud-native Data Security Posture Management (DSPM) solution continuously discovers and classifies all cloud data, structured and unstructured, across managed and self-hosted data stores, including unknown shadow data, without the data ever leaving your environment. It analyzes access, usage patterns, and security posture, and provides actionable, guided remediation for data security risk. Laminar connects to your multi-cloud environment including AWS, Azure, GCP, and Snowflake via APIs and is agentless, asynchronous, and completely autonomous.

Data Security Posture Management (DSPM)

Laminar Special Edition

by Lawrence Miller

for dummies®
A Wiley Brand

Data Security Posture Management (DSPM) For Dummies®, Laminar Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN 978-1-394-18106-3 (pbk); ISBN 978-1-394-18107-0 (ebk)

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Project Manager: Jen Bingham
Acquisitions Editor: Traci Martin
Editorial Manager: Rev Mengle
Sales Manager: Molly Daugherty
Content Refinement Specialist: Tamilmani Varadharaj

Table of Contents

Introduction

The universe is expanding. But we're not talking about planets, stars, and galaxies. The digital data universe — which has quite a few black holes of its own — is growing exponentially. According to earthweb.com, there are now more bytes of digital data than observable stars in the universe, and it's predicted that there will be about 200 zettabytes (roughly 186 trillion gigabytes!) of data by 2025.

Data security has long been top of mind for organizations everywhere. However, the proliferation of data — spurred on by digital transformation, cloud-first strategies, advanced analytics and machine learning, and DevOps application development pipelines — has accelerated the need for an automated, scalable, and more agile solution.

Data security posture management (DSPM) solutions address this need across the full data security life cycle — from discovery, classification, cataloging, and risk prioritization to access control, policy enforcement, remediation, and real-time monitoring.

Foolish Assumptions

It's been said that most assumptions have outlived their usefulness, but we assume a few things nonetheless!

Mainly, we assume that you are looking for a solution to address data security in the cloud. Whether you are a Chief Information Security Officer (CISO), a data security, privacy, or governance practitioner, or a security or cloud architect, this book will explain how DSPM can help you address your data security needs.

If any of these assumptions describe you, then this is the book for you! If none of these assumptions describe you, keep reading anyway. It's a great book, and after reading it you won't be posturing about your knowledge of data security posture management.

Icons Used in This Book

Throughout this book, we occasionally use special icons to call attention to important information. Here's what to expect:

REMEMBER

This icon points out important information you should commit to your nonvolatile memory, your gray matter, or your noggin.

TIP

Tips are appreciated, but never expected – and we sure hope you'll appreciate these useful nuggets of information.

WARNING

These alerts point out the stuff your mother warned you about (well, probably not), but they do offer practical advice.

Beyond the Book

There's only so much we can cover in this short book, so if you find yourself at the end of this book wondering, "Where can I learn more?" go to `https://www.laminarsecurity.com`.

IN THIS CHAPTER

- Exploring the ever-changing attack surface and threat landscape
- Getting a handle on data security challenges in the cloud
- Understanding the limitations of existing data security options

Chapter 1
Just How Secure Is Your Data in the Public Cloud?

This chapter explores modern trends that are defining the new attack surface, the challenges of data security in the cloud, and the limitations of many current approaches to data security.

How We Got Here from There: The New Innovation Attack Surface

For modern enterprises, data is at the center of innovation. These companies understand that data is a key asset and a source of competitive differentiation. They democratize data to unleash its full potential and make it accessible for application developers, data scientists, and business users. The issue is that as data proliferates, security doesn't travel with it — and adding the pace of change to the sprawl of cloud technology means that data security teams just can't keep up. Data is the fuel for innovation and, for modern organizations, data represents their "crown jewels." Malicious actors know this all too well and constantly target this new threat vector — the "innovation attack surface."

Most organizations unconsciously accept the innovation attack surface as the cost of doing business. They accept the continuous unintentional risk that cloud data users create when data proliferates, the source of the innovation attack surface. This new attack surface contrasts with traditional attack surfaces, which are determined by external forces (including bad internal actors) seeking to exploit vulnerabilities to gain illicit access to protected information.

Some important trends can help you understand how all of this has transpired.

Cloud transformation and data democratization

Today, innovation happens in the cloud. This cloud transformation is great for the business, but it also introduces new cybersecurity risks and requires changes to people, processes, and technology.

REMEMBER

Cloud transformation isn't new. It's been ramping up for over a decade. However, the COVID-19 pandemic drove companies to fast-track their plans to provide new or additional digital services to customers and employees.

The cloud has also enabled widespread data democratization, enabling easy access to essential data for developers, data scientists, and business users to support their innovation efforts. However, this freedom to access and use data without oversight creates unknown, unmanaged, and unprotected cloud data sources, resembling the chaotic Wild West.

WARNING

Cloud transformation and data democratization are creating the perfect storm, giving rise to the innovation attack surface.

Technology sprawl and complexity

In the public cloud, it's all about change and innovation. Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) are in a constant battle to roll out new services for their customers. Sounds great, right? Well, sometimes it can be too much of a good thing. Each cloud service is configured and used differently, and each introduces new and unique risks. The ever-changing architectures are confusing and complex, and if

you're not careful, this can lead to some costly and even devastating mistakes with sensitive data stored in the cloud.

Cloud data proliferation

Nearly half of all data (48 percent) is stored in the public cloud today and it's only increasing, according to the *Flexera 2022 State of the Cloud Report*. Cloud transformation and data democratization (discussed earlier in this chapter) further exacerbates the proliferation of data in the cloud. Unfortunately, traditional data security controls are unable to keep up with the dynamic movement of data, so they must be configured from scratch every time data is created, copied, shared, or moved.

WARNING

Developers and data scientists are major contributors to data proliferation and shadow data because they have permissions and technologies they can use to easily create or duplicate entire data stores. They can do this without oversight from security or compliance teams and without pressure to delete the data after it's no longer needed.

Death of the traditional perimeter

One of the many benefits of the cloud is that it is accessible from anywhere. Remote users can access their applications and data whether working in the office, at home, or in a hotel lobby. The notion of a network perimeter — an on-premises data center protected by a firewall — has all but disappeared. The lack of a single choke point (a firewall) means sensitive data is exposed by design because anyone can access it from anywhere with the proper credentials (whether authorized or stolen).

TIP

To protect sensitive data in the modern cloud era, security controls must move from the perimeter to (and with) the workloads and data elements themselves.

Faster rate of change

Competitive pressures are unrelenting in a global economy that thrives on innovation and faster time-to-market. As a result, release cycles now happen in weeks, days, and hours rather than months and years. Unfortunately, security teams are usually not on that same quick schedule and still rely on slower manual approaches. This highlights the need for agile, automated solutions to match the speed of innovation.

The changing role of security

Cloud transformation and data democratization offer significant advantages for businesses, but data security teams must evolve to securely enable the business rather than just slowing everyone down or letting risk grow exponentially. Data security must focus on protecting data from breaches and compromises while also empowering users to be productive. To achieve this, data security requires an agile and always-on mindset.

What's So Challenging About Data Security in Public Cloud?

Cloud service providers use the shared responsibility model (see Figure 1-1) to delineate their and their customers' responsibilities for operating, maintaining, and securing different elements of the cloud stack. This model takes some burden away from customers, depending on the services they use, but regardless of the services, customers are always responsible for their data.

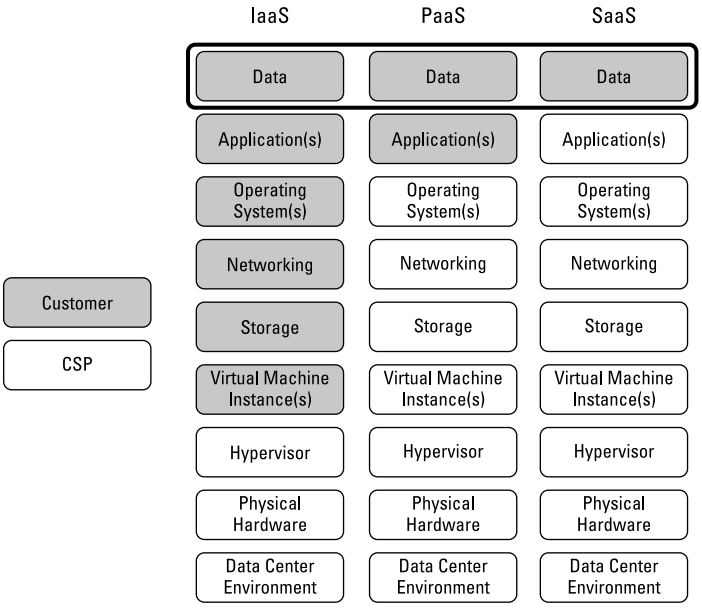


FIGURE 1-1: The shared responsibility model shows who is responsible for what in the public cloud.

With this important responsibility in mind, here are a few of the reasons data security is so challenging in the cloud:

- **Lack of data visibility and control:** Security teams struggle to discover, identify, manage, and control sensitive data that proliferates across hybrid and multicloud environments.
- **Rapid rate of change:** As discussed earlier in this chapter, change is constant in the cloud. This is true of both the technologies and services offered in the cloud, as well as the data itself in the cloud. Staying on top of these changes is a constant struggle for security teams.
- **Unrestricted access and use of data:** For authorized users, access to data in the cloud is pretty much unrestricted. They can access data, copy it, modify it, share it, and move it as needed — which means more data proliferation and change.
- **Cloud security resources and skills gap:** The shortage of cybersecurity professionals globally is compounded in cloud security due to differences in public cloud offerings and constant changes. According to ISC2, the global cybersecurity workforce has a shortage of 3.4 million professionals, with cloud data security expertise even scarcer. While cloud infrastructure security is increasingly popular, understanding cloud data security remains a shared responsibility, and resources with such knowledge are limited.

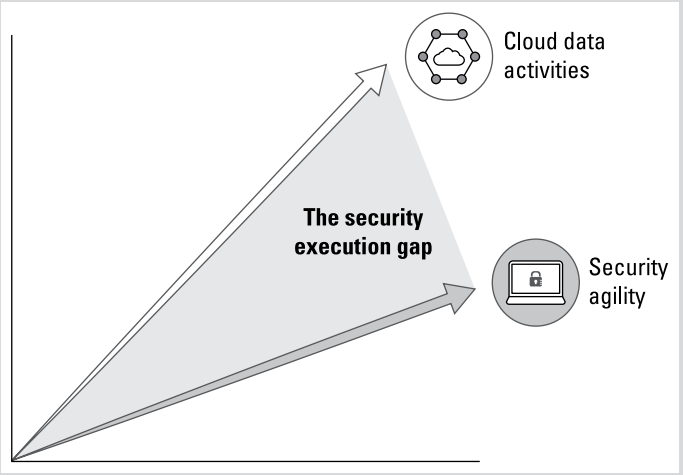
CLOSING THE DATA SECURITY EXECUTION GAP

The introduction of the innovation attack surface, as discussed earlier in this chapter, has given rise to a new problem: the security execution gap (see the nearby figure). This gap is best described as the growing divide between the activities that contribute to business innovation and the security activities and expertise necessary to safeguard the business. Until recently, marrying organizational agility and innovation with the necessary security controls was impossible. With the exponential proliferation of data in the cloud, finding automated solutions to protect cloud data is critical in closing this gap. Chapter 3 details a DSPM solution that handles the complexities resulting from the sprawl of technology and easily guides the organization to an

(continued)

(continued)

improved security posture. This ultimately eliminates manual efforts and augments the level of security expertise, which can help optimize the role of data security experts on the team.



Bumping into the Limits of Existing Solutions

Security teams face unprecedented challenges protecting sensitive data in the cloud due to the lack of proper solutions. Limited resources and outdated manual processes only add to the difficulty, leaving them blind to sensitive data in the cloud and struggling to keep up with increasing risk.

Don't do anything

Ignoring your data security challenges in the cloud is not a viable solution. This ostrich approach of burying your head in the sand and hoping for the best can lead to regulatory fines, lawsuits, and even bankruptcy. Plus, it's not a long-term solution, as the problem will only worsen as cloud data continues to proliferate. "Security through obscurity" is not an option. While your

unknown data may be obscured to you in the cloud, it's practically screaming "pick me, pick me" to malicious actors.

Do it yourself

While some organizations may choose a DIY approach to data security for the cloud, it's important to realize that it's not as simple as a typical home improvement project. Building and maintaining a custom solution can ultimately cost organizations more in the long run. There are several challenges that arise when attempting to take on this task, including difficulty keeping up with dynamic public cloud environments, incomplete security coverage, and limited understanding of data exposure. These factors make it difficult to fully understand and develop solutions for the many cloud services available, leaving unfound shadow data still at risk or missing the valuable operationalization of posture management.

Legacy (on-premises) data security solutions

Organizations that initially used lift-and-shift migrations of their on-premises systems and applications to move to the cloud may want to consider adopting more modern cloud-native solutions, including for data security. A lift-and-shift migration of legacy on-premises data security solutions will fail due to their legacy, connector-based architecture. Needing to connect to each data asset leads to difficulty of configuration and maintenance, the inability to autonomously discover unknown or offline data, the need for credentials to access each data asset, and a long time to value. As a result, security teams will struggle to keep up with the thousands of data assets that change daily in a dynamic cloud environment.

Take an example of an organization with over 1,000 EBS instances. Each one needs to be connected to manually, including finding access credentials. The initial configuration can take months, let alone the constant maintenance as change occurs. The operationalization cost quickly becomes unbearable and higher than the value.

CSP-native tools

Although major public clouds like AWS, Azure, and GCP offer their own CSP-native tools to help with cloud data security, these tools have some limitations to be aware of. They are often designed and optimized for specific clouds, making them single-cloud specific or offering limited multicloud support. Additionally, they lack continuous scanning capability, making them cost-prohibitive for many organizations. Furthermore, they have a limited data store scope, such as AWS Macie supporting only AWS S3 buckets, and provide no posture management functionality. Finally, these tools only show data on assets that are configured for scanning, leaving shadow data unaccounted for.

Cloud infrastructure security solutions

Finally, a number of cloud infrastructure security solutions have emerged in recent years, such as cloud security posture management (CSPM) and cloud-native application protection platforms (CNAPPs). However, these solutions focus on the infrastructure layer and don't address data security in the cloud. For example, they don't identify which data should be encrypted, how long it should be retained, or who should have access to it. Additionally, they don't monitor access to sensitive data in the cloud or identify any indications of data leakage or exfiltration. These infrastructure-centric solutions serve a different user than pure-play DSPM. Data security teams care about the data, regardless of the underlying infrastructure it resides on.

REMEMBER

Some CSPM and CNAPP vendors are now including basic data security capabilities in their solutions — but they're not the same thing as DSPM. CSPM, CNAPP, and DSPM all serve different users and use cases. CSPM and CNAPP solutions focus on infrastructure (and application) security, with little concern for data governance and privacy requirements. DSPM solutions, on the other hand, focus exclusively on data security, governance, and privacy.

In Laminar's 2023 State of the Public Cloud Security Market Report respondents answered that 97 percent of organizations now have a dedicated team for data security. If you don't, you should.

IN THIS CHAPTER

- Defining shadow data
- Recognizing the shadow data problem
- Exposing the many places that shadow data can hide
- Understanding how sensitive data can be compromised

Chapter 2
Shining a Light on Cloud Data Risk

Data tends to proliferate in the public cloud. There often seems to be a set-it-and-forget-it mentality when it comes to the cloud. In fact, the public cloud is sometimes described as just being someone else's data center, but that doesn't mean the security of your data in the public cloud is someone else's problem. Data is a powerful enabler of innovation, and for many modern enterprises it is their most valuable asset. However, the data you don't know about — lurking in the shadows — could come back to haunt you.

This chapter discusses shadow data: what it is, why it matters, where it hides, and how it exposes your organization to risk.

Shadow Data Is the New Shadow IT

Shadow IT consists of departments and users leveraging technology (such as personal devices and unsanctioned applications) that isn't authorized or managed by an organization's IT department.

Shadow IT creates risk for organizations, including unknown security vulnerabilities, regulatory noncompliance, lack of support, failure of business-critical processes, and procurement inefficiencies and waste.

Like shadow IT, shadow data is data that is created by different users throughout the organization — such as application developers and data scientists — but IT and security teams aren't aware of shadow data, and it often goes unmanaged and underprotected (see Figure 2-1).

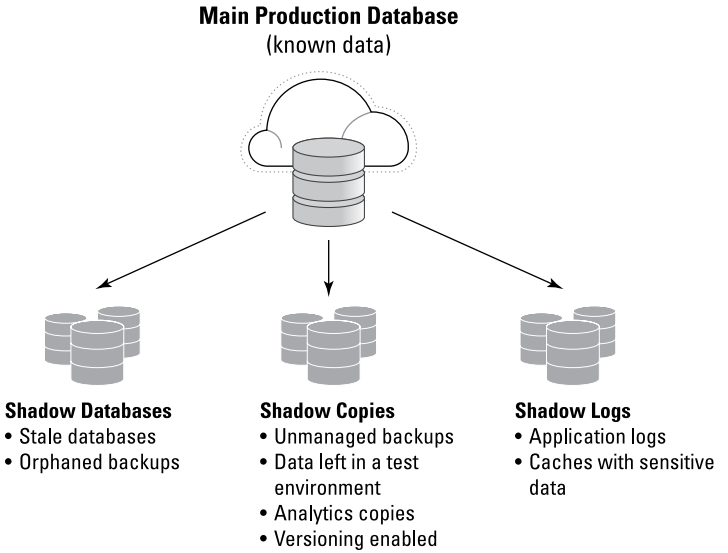


FIGURE 2-1: Shadow data is created as users make copies from a production database not governed by or known to security.


REMEMBER

Like shadow IT, we all have shadow data, but we don't know how bad it is. Shadow data is any data that is not governed, has no oversight, and typically is less protected. In other words, data that is unmanaged and unknown, which means your security team is effectively blind to it.

For example, DevOps teams and data scientists may spin up a data store in the public cloud to copy sensitive production data for use in a development or sandbox environment, or to train machine learning (ML) models. IT and security typically will be totally unaware of the additional shadow copy or copies.

SHADOW DATA BREACH EXAMPLE: COMMUTEAIR

In January 2023, a 2019 copy of the U.S. Transportation Security Administration's (TSA) No Fly List was exfiltrated from CommuteAir's Amazon Web Services (AWS) environment. The threat actor was able to steal CommuteAir's AWS credentials from an underprotected development server, leading to the compromise of more than 1.5 million records in the No Fly List and more than 250,000 records in the Secondary Security Screening Selection (SSSS) List. Both lists contain sensitive personally identifiable information (PII) including individual's names and dates of birth. Additionally, a database containing PII of CommuteAir employees was compromised.

According to CommuteAir, the lists were used for testing software-based compliance processes for implementing federally-mandated security requirements.

This breach could have been avoided if data protection teams were aware of the existence of this sensitive data and the security posture of the asset. However, you can't protect what you can't see, and today data protection teams are blind to shadow data in public cloud environments, even if it includes sensitive data.

TIP

According to the Laminar *State of Public Cloud Data Security Report 2023*, 93 percent of security practitioners are concerned about shadow data and it has emerged as the No. 1 challenge in cloud security.

Why Is Shadow Data a Problem?

Like shadow IT, shadow data — caused by data proliferation — is a growing problem because it increases organizational risk. These risks include security issues, data privacy and governance challenges, and compliance violations.

REMEMBER

Data will continue to proliferate. This is the reality of the cloud, and data security needs to become agile to match.

From a security perspective, shadow data may reside in public cloud data stores that are not protected by standard security controls and may also be misconfigured. This is a common problem in AWS S3 buckets and Microsoft Azure blob storage accounts — at one time, these storage services were open to the public by default!

TIP

Although most public cloud providers have since addressed some of these risky default configurations, you should check all of your public cloud storage resources in case any were created before the defaults were changed. There are also still many security features that are not, but should be, configured by default.

Shadow data also creates many compliance issues for organizations. Cardholder data environments (CDEs) are subject to the Payment Card Industry (PCI) Data Security Standard (DSS). If a developer makes a copy of the data in the CDE that isn't properly masked or anonymized, the data can be breached — leading to significant penalties for the company. Similarly, the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), among others, grant an individual the "right to be forgotten" — that is, to have all of their private data deleted. If you have shadow copies of private data, you again risk substantial penalties due to noncompliance. Some data sovereignty requirements restrict where certain sensitive data can be stored to a specific country or region. These requirements may be violated by the existence of shadow data in unauthorized locations.

Data backup and recovery are also often overlooked with regard to shadow data. On the one hand, shadow data isn't known or managed by the organization, so you might take the attitude of "use shadow data at your own risk." However, the shadow data was presumably created for a legitimate business purpose and the organization is still responsible for the data. Losing a massive data store, for example, that is used by a data scientist to train an ML model for a new project can cost the organization quite a bit.

Finally, although storage in the public cloud is relatively inexpensive compared to traditional on-premises storage infrastructure, there is still a cost. And that cost can grow quickly when you have multiple copies of the same data in multiple clouds and data stores or unused copies that should have been deleted but were not.

Where Can Shadow Data Hide?

Shadow data is everywhere. Unfortunately, it's not like playing a game of hide-and-seek where you only have to find one hiding place to win. Instead, you have to find every hiding place!

Some common scenarios that lead to the creation of shadow data in an organization and increase risk — especially if it is later forgotten — include the following:

- An application developer creates a database copy to run a test in a development environment and later forgets to delete the copy.
- A structured query language (SQL) database is moved to the cloud for a lift-and-shift project. The database then gets refactored into a cloud-managed relational database service, but the original SQL database is never removed.
- A legacy application is decommissioned but the backup database that was associated with the defunct application remains.
- A developer spins up an embedded database inside a compute instance but doesn't notify security. To IT and security, this database looks like just another compute instance, and they don't realize it is actually a data storage source.

Table 2-1 shows a few more examples of shadow data and where it can hide in the public cloud.

TABLE 2-1 Shadow Data in Public Cloud Environments

Shadow Data Type	Created By	Example
Database copy in test environment	Developers	Developers create copies of RDS databases (for example, by using a dumping utility such as pg_dump). These copies create unmanaged datastores with sensitive user data, without the proper guardrails.
Unmanaged backups	DevOps	DevOps backs up data to storage buckets, but it is not documented. These buckets create large amounts of data that can easily be exposed to the public.

(continued)

TABLE 2-1 *(continued)*

Shadow Data Type	Created By	Example
Toxic application logs and cache	Developers	Developers and log frameworks log sensitive data, which creates sensitive files that are not classified as sensitive, may be easily exposed, and lack proper access controls and encryption.
Analytics pipeline	Data scientists	Data scientists take copies of databases and data from a data lake and move it using extract-transform-load (ETL) tools that can either be hosted, managed, or software as a service (SaaS) based. These tools copy the data to uncontrolled locations that are not managed by the proper guardrails and may expose the data.
Stale unmaintained databases	Cloud architects	Data architecture gets changed rapidly due to business needs. Old data stores left over from lift-and-shift projects are kept for recovery, but are often forgotten long after the migration process is completed.
Unlisted embedded databases	Developers	Easy to spin up, but hard to detect embedded databases on compute resources such as PostgreSQL in an AWS EC2 instance can be hard to distinguish from other EC2 instances, but equally as risky as RDS.

What Are the Different Ways Sensitive Data Can Be Exposed?

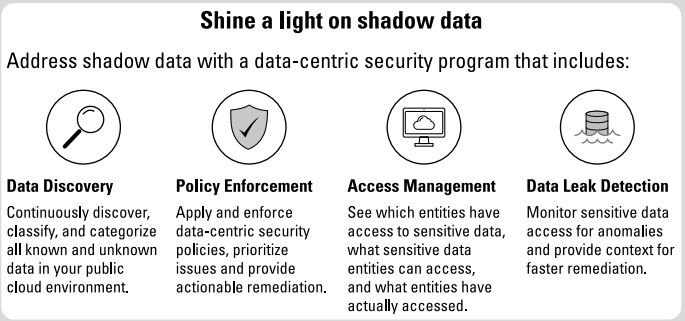
As the volume of shadow data increases, so too does your risk of a data breach. Shadow data stores are more likely to be miscon-figured, unmonitored, and violate your data security policies — making them easy targets for malicious actors.

In the rush to deliver innovative new products to market, application developers may overlook important data security policies — such as controlled access, retention, backups, encryption, and anonymization — potentially exposing the organization to a data breach.

WHAT'S LURKING IN THE SHADOWS? SOME EXAMPLES OF WHAT CAN BE FOUND IN SHADOW DATA

Any data can be shadow data, but some data is more sensitive than the rest. Here are some common examples of sensitive data that needs to be protected:

- **Personally Identifiable Information (PII)** such as social security numbers, driver's license numbers, passport numbers, and so on
- **Financial information** such as credit card numbers, bank account numbers, cardholder data, credit scores, and so on
- **Protected Health Information (PHI)** such as medical history, health data, lab results, prescribed medications, and so on
- **Intellectual property (IP)** such as trade secrets
- **Account credentials** such as usernames and passwords



Insider risk is another potential source of exposure. Insider risk can exist due to misconfiguration of public cloud data stores (including sharing permissions), accidental sharing of sensitive data, or malicious exfiltration of data (for example, just prior to quitting a job).

TIP

In Chapter 3, you'll discover how to address shadow data and other data security risks in the cloud with cloud-native DSPM.

IN THIS CHAPTER

- What should a DSPM actually do?
- Putting your DSPM solution to good use (cases)

Chapter 3
Putting First Things First With a Data-First Approach

Data is the lifeblood of modern businesses, powering crucial insights, informing decisions, and driving improved outcomes. One significant challenge for data-driven organizations is efficiently enabling broad access to valuable data while ensuring its security and protection. Adopting a data-first, or datacentric, approach to security — rather than an infrastructure-centric one — helps organizations tackle this challenge.

This chapter explores how a successful data security strategy for the cloud begins with a data security posture management (DSPM) solution and outlines various use cases that a DSPM solution can address for organizations.

The Importance of a DSPM Solution

A datacentric approach emphasizes the importance of protecting your valuable data rather than focusing on systems and infrastructure. After all, while attackers may target your systems and infrastructure, their ultimate goal more often than not is to steal your valuable data — a burglar may break open a safe, but his ultimate goal is to steal what's inside the safe.

A DSPM solution empowers organizations to implement a datacentric security strategy by first providing an accurate inventory of their sensitive data and identifying where it violates data security policy, thereby enhancing overall data security posture.

According to a Gartner report, Gartner defines data security posture management (DSPM) as providing "visibility as to where sensitive data is, who has access to that data, how it has been used, and what the security posture of the data store or application is."

With this in mind, let's dig into the key capabilities of a DSPM solution, which include global data visibility, data hygiene, data security risk control, data access governance, and privacy and compliance.

Global data visibility

Global data visibility provides organizations with a comprehensive view of their sensitive data. After all, you need to know what you're safeguarding to ensure it is appropriately protected. This involves identifying the location and type of sensitive data to ensure proper protection measures are in place. To achieve a global view, all clouds — including infrastructure-as-a-service (IaaS), platform-as-service (PaaS), and software-as-a-service (SaaS) resources — need to be covered. The appropriate data owners must also be determined to facilitate efficient communication of any data-related security or privacy issues. In larger organizations, this global view must be further broken down by business units or teams so proper ownership can be operationalized.

Data hygiene

Data hygiene, as you might imagine, is about keeping your data clean and healthy. It encompasses various actions that help organizations maintain clean and organized data in accordance with their data governance framework. This includes addressing and

remediating misplaced, redundant, and obsolete data to streamline maintenance, optimizing storage resources, and reducing potential security risks. Purging outdated or irrelevant data is another essential part of good data hygiene, resulting in the retention of only accurate and useful data. Data hygiene is an ongoing effort (much like brushing your teeth and washing dirty dishes isn't a "one-and-done" effort), and as such, it should be done continuously through established policies that set guardrails for maintaining the overall quality, usefulness, and security of the data.


TIP

Purging outdated and irrelevant data isn't just good for hygiene. It can also reduce your legal liability in the event of a data breach. According to the Ponemon Institute's *Cost of a Data Breach Report 2022*, the cost of a data breach in 2022 was $164 per record. This means that a company that, for example, purges one million outdated or irrelevant data records containing personally identifiable information (PII) or other regulated data, can avoid more than $160 million of potential liability in the event of a breach.

Data security risk control

Data security risk control involves immediately detecting and proactively remediating data risk factors to prevent data breaches. This capability detects and addresses three key data postures:

- **Overexposed data,** such as public read access, or permissive access rights which should be identified and mitigated to reduce the likelihood of unauthorized access or data breaches
- **Underprotected data,** where there are missing controls like encryption, masking, or proper retention policies
- **Misplaced data,** such as cardholder data subject to the Payment Card Industry Data Security Standards (PCI DSS) in an unauthorized environment or PII data in a development environment


REMEMBER

When performed correctly, security issues are prioritized based on risk to achieve the most effective remediation for expended resources.

Data access governance

Data access governance manages and controls access to sensitive data. This involves identifying all internal and external users,

roles, and resources with access to sensitive data, monitoring and controlling access patterns based on their roles and responsibilities, ensuring that only authorized users have access to sensitive assets, and regularly reviewing and updating access permissions based on actual usage. Particular attention should be paid to third-party access to data as it is typically less closely governed and often involves shared access accounts.

Privacy and compliance

Privacy and compliance ensure that organizations adhere to data privacy regulations and industry standards, and make audits more manageable (and perhaps a little less painful and costly). Providing objective evidence for audits can be challenging, but having reporting from your DSPM that shows you know where your data is and understand that its security posture can significantly ease compliance. DSPM allows governance, risk, and compliance (GRC) practitioners to define policies for automated validation and control. A DSPM policy engine should include built-in policies that adhere to data privacy and compliance frameworks.

REMEMBER

DSPM serves as the technological cornerstone of a datacentric security strategy. However, to effectively implement data security in the cloud, it's essential to balance people, processes, and technology. Enterprise data responsibility involves multiple teams, including data security, privacy, development, DevOps, and governance. DSPM can assist these cross-functional teams in streamlining data security by integrating smoothly with existing business processes and technologies.

A DAY IN THE LIFE CYCLE OF DSPM

Data Security Posture Management is not a linear one-and-done activity. Rather, it's a continuous cycle (the D-P-S-M of DSPM) that spans four broad activities:

- **Discover:** Continuously and autonomously discover, classify, and catalog all known and shadow data at scale across your cloud environments. Assess the security posture status of sensitive data against datacentric security policies and compliance requirements.

- **Prioritize:** Prioritize all issues according to their risk profile based on sensitivity level, security posture, volume, and exposure.
- **Secure:** Alert on policy violations, detail clear and concise explanations, including object-level evidence, then provide guided and actionable remediation recommendations.
- **Monitor:** Continuously monitor new and modified data assets against stated security posture and compliance requirements throughout its life cycle, regardless of where the data moves in the cloud.

What Can You Do With DSPM? Common DSPM Use Cases

Data security, governance, and privacy teams can use DSPM solutions in many ways to help keep their organization secure and compliant. Here are a few of the most common use cases for DSPM.

Automating data discovery and classification

In the cloud, with its proliferation of data and rapid pace of change, creating and maintaining an accurate data inventory is virtually impossible without automation.

DSPM helps organizations automatically and continuously discover, classify, and categorize all their known and unknown data — including sensitive, proprietary, regulated, abandoned, and shadow data — across multicloud environments such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), Snowflake, Office 365, and more.

DSPM autonomously builds and continuously updates an extensive and easy-to-consume data mapping to meet your security, privacy, and governance requirements. You can stay up to date and track changes without having to rescan unchanged data on your own. There's no need for connectors, access credentials, or lists of data assets, and most importantly, the data should never be sent outside your environment.

Enforcing data security policies automatically

You may already have detailed data security policies, but how do you know they are being followed? If you're using manual efforts, then your task is hopeless.

DPSM automatically enforces data security policies at scale for all your data as it travels through the cloud. DSPM converts data policies into specific technical configurations and shows where data security policies are violated, prioritizes issues for resolution, and helps you fix those issues with clear, specific technical remediation instructions.

With DSPM, you can:

- Identify your most valuable data that must be protected, considering the data's content and current security posture.
- Continuously monitor security posture drift against extensive prebuilt policies aligned with security best practices, governance frameworks, and privacy regulations.
- Customize existing policies or create your own to meet your unique business requirements.
- Get notified (or alerted) when violations occur and receive actionable remediation advice via seamless integrations with your existing ticketing and workflow tools.

TIP

DSPM datacentric policies focus on violations such as:

- **Overexposed data:** For example, public read access, excessive user access, third-party access, and unused internal access
- **Unprotected data:** For example, missing encryption at rest, masking, encryption in transit, plaintext credentials, activity logging, and retention
- **Misplaced data:** For example, European Union citizen data outside of General Data Protection Regulation (GDPR) countries, personal, medical, and financial data in lower (such as development, test, and staging) environments
- **Redundant data:** For example, unmanaged backups and sensitive data in abandoned assets

Controlling data exposure

As data rapidly proliferates in the cloud, security does not follow that data, often leading to crucial business data being exposed.

DSPM pinpoints all your exposed sensitive data that can lead to adverse outcomes like data breaches, ransomware attacks, and noncompliance penalties — whether it's misplaced data (for example, sensitive data mistakenly stored in public buckets), misconfigured controls (for example, third-party access granted to sensitive data), or overly permissive access. You can use out-of-the-box or custom policies to automatically detect, prioritize, and monitor exposure violations and remediate them with detailed evidence and actionable, guided recommendations.

Controlling datacentric environment segmentation

Many organizations segment their environment to address unique business needs and comply with various data security and privacy requirements such as PCI DSS, the U.S. Health Insurance Portability and Accountability Act (HIPAA), and the European Union General Data Protection Regulation (GDPR). However, manually implementing and enforcing segmentation is impossible in dynamic cloud environments.

DSPM helps you segment your cloud environments and apply location controls to comply with security and regulatory requirements. You can detect and receive alerts when sensitive or regulated data is placed in untrusted and/or unauthorized environments, review violations, and take action to remove the data or authorize the new environment.

Complying with data privacy and compliance frameworks

Regulatory compliance is a complex and daunting challenge, but proper evidence for audits can make things much easier (and less costly).

DSPM streamlines evidence collection for internal and external privacy and governance stakeholders through autonomous data discovery and classification of your sensitive and regulated data. A DSPM data policy engine continuously enforces regulatory compliance and standards requirements for data, regardless of the underlying technology or location. You can easily ensure that all data is properly owned and tagged to fast-track evidence collection for Records of Processing Activities (RopA) and to answer Data Subject Access Requests (DSAR).

Chapter 4
Ten Must-Haves for a DSPM Solution

This chapter discusses ten must-have requirements to look for in a data security posture management (DSPM) solution for your organization.

- **Autonomous:** Rather than just inspecting known assets, your DSPM solution must be able to automatically discover unknown, new, and modified data stores across all of your clouds — without needing credentials or manual configuration.
- **Continuous:** Change is constant, especially in the public cloud. To keep up, your DSPM solution must be able to continuously monitor your environment for changes and automatically scan new cloud accounts, new data stores, and new data added to existing data stores.
- **Secure by design:** When evaluating a DSPM solution, the last thing you need is another source of risk. Look for a solution that doesn't extract data from your environment — your DSPM should use the cloud service provider's (CSP) application programming interface (API) and ephemeral serverless functions in your cloud account to scan your data.
- **Breadth and depth of coverage:** If you're like most organizations, you probably have a multicloud environment — so you need a DSPM solution that provides consistent security and governance across all your clouds. Whether you're using Amazon Web Services (AWS), Azure, Google Cloud Platform (GCP), Snowflake, Office 365, or practically any combination of various

cloud database, storage services, or software as a service (SaaS) apps, you need a single and consistent view of your data across clouds, geographies, and organizational boundaries to evaluate the risk to your data across all clouds. Your DSPM solution also needs to support variations of technologies in each data asset subset, including both managed and self-hosted databases with unique configurations and many file types covering structured and unstructured data.

- **Intelligent classification:** Look for a solution that utilizes multistep contextual analysis to automatically identify sensitive data with low false positive (FP) and false negative (FN) rates. The solution should also include hundreds of predefined classification rules, data validators, and classification algorithms that extract the data insights you need without having to locate the data owner.
- **Extensive set of built-in datacentric policies:** Don't reinvent the wheel. Instead, look for a solution that provides out-of-the box datacentric policies for common use cases like data security, proper governance, and privacy. Policies cover data security requirements like encryption; retention, archiving, and environment segmentation; and guidance on who is allowed to access what data; as well as security policies that cover overexposed, underprotected, and misplaced data.
- **Customization features:** You need a solution with robust customization features that are flexible and powerful enough to match your data taxonomy and address any unique requirements your organization may have such as sensitivity levels/definitions, data types, and custom industry policies.
- **Guided remediation:** Look for a solution that provides a full analysis of why a security or compliance violation exists, evidence of its existence, and technical recommendations on how to fix it based on policy and environment.
- **Simple and quick deployment process:** Your DSPM solution should be agentless and connectorless to simplify and accelerate the deployment process. Look for a solution that can be deployed in minutes and delivers time-to-value in a few days.
- **Easy integration with your ecosystem:** Look for a DSPM solution with extensive integrations that include third-party systems such as IT service management (ITSM), security information and event management (SIEM), cloud security posture management (CSPM), extended detection and response (XDR), and data catalogs.

Laminar

Your Data Security in the Cloud Strategy Continues...

Read:
'A Buyer's Guide to DSPM Solutions'

Find out more about data security in the cloud

Today, data is the lifeblood of modern business and keeping data protected in the cloud is key. As companies innovate in the cloud, data proliferates at an extraordinary pace. Data security posture management (DSPM) solutions address this challenge across the full data lifecycle. This helpful book helps you understand exactly what you need to look for in a DSPM solution

Inside...

- Helpful use cases
- Challenges with securing data in the cloud
- Getting to know shadow data
- How you can keep data secure in the cloud
- Three core criteria and ten must haves for a DSPM solution

Lawrence Miller is an information technology professional and the author of more than 180 Dummies books.

Go to Dummies.com™
for videos, step-by-step photos, how-to articles, or to shop!

ISBN: 978-1-394-18106-3
Not For Resale

for dummies®
A Wiley Brand

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.