SailPoint

WHITE PAPER

Securing Access to Files **with Identity Governance**

What IAM Leaders Need to Know

Throughout most enterprises, identity and access management (IAM) directors and project leaders face a growing security risk represented by sensitive data stored in unsecured files, also referred to as unstructured data. This paper provides practical advice on how IAM managers should respond to this risk and explains how identity governance can be extended to better secure unstructured data to meet privacy and compliance requirements.

Controlling access to unstructured data is a growing problem for organizations around the world. Most enterprises have far more unstructured data than they realize, and that data is growing at an accelerated rate. According to Gartner, upward of 80% of enterprise data today is unstructured, yet many do not have adequate visibility and control over their unstructured data.

What is Unstructured Data?

In its most basic definition, unstructured data simply means any form of data that does not easily fit into a relational model or a set of database tables. Unstructured data exists in a variety of forms, including documents, spreadsheets, presentations, and reports, and is typically stored in individual files.

The risks associated with unstructured data are significant. The vast majority of unstructured data is created, extracted or downloaded by individuals, and is stored and shared from a variety of locations, often outside the purview of the IT department. While IT staff may believe mission-critical data in structured format is secured, it is a fairly simple process to extract highly sensitive data from ERP and mainframe systems in order to create shareable files.

The Growing Risk of Unstructured Data

Most organizations have not adequately safeguarded the information assets and sensitive data stored in unstructured files, mainly because they do not know how much they have or where it resides. As a result, they are at serious risk of security breaches and compliance penalties.

Professional hackers are increasingly targeting unstructured data because it is often easier to steal and yields a treasure trove of valuable information. In the past year, there were dozens of examples of data breaches involving the theft of email archives, legal contracts, medical documents, customer lists, trade secrets, source code, and other highly sensitive material – all stored in files as unstructured data.

CASE IN POINT 1

DNC

In September 2016, former Secretary of State Colin Powell's personal emails were stolen from the Democratic National Committee and publicly posted on DCLeaks.com. Included in the emails was a highly confidential attachment listing Salesforce's acquisition targets and the details of its M&A strategy. (Powell is a Salesforce board member.)

CASE IN POINT 2

Gorilla Glue

In November 2016, hackers stole 500GB of data from adhesive manufacturer Gorilla Glue. The data included research reports, product designs, strategy documents, and financial spreadsheets. Based on prior crimes committed by the hackers, it is believed they demanded a ransom payment from Gorilla Glue.

Failure to secure sensitive data stored in files not only increases the risk of a data breach, it also increases regulatory risk. New privacy regulations, such as the European Union General Data Protection Regulation (GDPR), have introduced new and stringent requirements for handling personal data and established harsh penalties for failure to adequately secure it. In general, privacy mandates such as GDPR and the Healthcare Insurance Portability and Accountability Act (HIPAA) require appropriate security over both structured and unstructured data. Historically, organizations have focused most of their efforts on protecting structured data, but failure to address unstructured data could result in severe fines and legal penalties.

Why Is Unstructured Data Not Being Secured?

There are many reasons why organizations currently fall short on addressing the security and privacy risks associated with unstructured data. Let's examine each of the organizational and technical barriers to success in greater detail.

1 Failure to Adapt to the "New Normal"

Most organizations do not understand how much sensitive data they have stored in unstructured formats. Analysts estimate that the amount of unstructured data greatly exceeds structured data in most enterprises. This is a big shift from a decade ago, and many organizations have not evolved their security and compliance efforts to the "new normal."

Very often, there is a giant mismatch between the amount of unstructured data that must be managed and the investment in staff and tools to manage it. Most organizations still spend the lion's share of their security budgets on mission-critical applications, databases and the platforms they run on, both in the cloud and in the datacenter. Likewise, auditors focus their controls and compliance testing on the same mission-critical resources.

Without a doubt, the budgets and resources spent on structured data must be maintained. The question is: how much additional budget is needed to mitigate the security and regulatory risk of unstructured data? And what is the best way to extend the capabilities of current security and compliance efforts?

2 Confusion About Data Ownership; "Whose Job Is It?"

In many organizations, it is not always clear which department owns the unstructured data problem. Business users create the majority of human-generated unstructured data, but most business groups do not feel responsibility for protecting unstructured data and may not even be aware of the issue.

IT ownership of unstructured data is also hard to pin down. There are several groups CIOs may ask to solve the problem. Common choices are the storage management group, the data security team, and in some cases, the IAM team. Each of these teams is likely to approach the problem differently, with different tools. As an IAM leader, it is important that you consider the organizational confusion and overlap that may occur:

- The storage management team owns the processes by which enterprise data is collected, shared, protected, cleaned and stored. When the storage team takes on the issue of unstructured data, it is usually with the goal to reduce storage costs and slow data proliferation. They may embrace data inventory or classification as a means to identify redundant or stale data. While this approach may reduce organizational risk by reducing the amount of unstructured data

that could be breached or exposed, it does not take into account the identity and business context of the data that identity governance can provide. Without this context, data can be improperly classified or purged, leading to compliance and security issues.

- The data security team is responsible for detecting and preventing internal and external threats. They employ a variety of tools for this purpose, including encryption, firewalls, security information and event monitoring (SIEM) and data loss protection (DLP) tools. While some SIEM and DLP products address the management of unstructured data, they do not provide full-featured identity governance capabilities. These solutions often ignore the fundamental access permissions layer, and lack full visibility or audit capabilities into who has access to data and what data is potentially overexposed. By complementing SIEM and DLP tools with the insight and control provided by a purpose-built identity governance platform, organizations can mitigate risks stemming from insider breach and other cyber threats.

What's missing when access to unstructured data is left out of the identity governance program? In short, organizations are forfeiting the tools, expertise, and established processes for ensuring rigorous control over granting and revoking access to high-risk data throughout a user's lifecycle with the enterprise. What's needed to adequately secure unstructured data are critical identity governance capabilities that ensure regular oversight over who has access to what, enforce access policy and remediate inappropriate access.

3 The Unmanageable Nature of Unstructured Data

From a management perspective, there are big differences between structured and unstructured data, all of which make unstructured data a giant challenge to manage and secure.

Because of the sheer volume of data – billions of files, and TBs or PBs of data – and the fact that it can be stored in a complex variety of locations, managing it requires extended capabilities that are complementary to identity governance functions. First, you will need tools to help you find and classify sensitive, high-risk data in order to prioritize the data you need to protect with strong access controls and governance processes. It is simply not practical to govern all unstructured data in the organization, so data classification is paramount.

Secondly, organizations will need a process for identifying data owners, which is extremely complex in the case of unstructured data stored in files. The majority of this content is created, shared, and stored by individual users. Sometimes it is possible to identify a departmental business owner or administrator, but oftentimes, extensive discovery and analysis is required to find and capture owners for unstructured data. In order to effectively govern the data, your identity governance team will need data owners that understand the nature of the data, who should have access to it and its potential risk to the business.

Does this scenario sound familiar?

As part of its Microsoft Office 365 initiative, a Fortune 500 company made the decision to migrate file stores to cloud repositories. When the IT Director took a closer look at the requirements, he realized how little visibility he had to what was stored in the petabytes of files that the company had accumulated over the years. Were all these documents still relevant? Did they contain sensitive information? Who should have access to these files and folders? Cleaning up sensitive data and how it is secured is paramount to any cloud migration program.

The third capability organizations will need is specialized entitlement analysis capabilities. Access control models vary dramatically across the different types of repositories for storing unstructured data. Understanding who can access a given file or set of files requires a complex analysis of permissions and groups on directory servers or other entitlement models. Once the identity governance team has the results of this analysis, it can be incorporated into existing processes for policy scans, access reviews and access requests.

4 Proliferation of Management Silos

Given the distributed nature of files that contain unstructured data and the lack of centralized ownership, it's tempting for individual groups within an organization to seek a tactical solution for managing it. In this scenario, each domain team implements its favorite tool to manage its part of the environment (e.g., Windows, SharePoint, UNIX/Linux or cloud), solving a short-term issue with a "point" solution. Most point solutions only cover a subset of unstructured data repositories, and they do not address structured data at all. The result is management silos.

Looking at the big picture, management silos have some major drawbacks. No one can see the overall state of affairs. Instead, IT is left juggling a number of tools that only provide visibility into one piece of the unstructured data environment. And management silos make it extremely difficult to apply a uniform set of security policies for access control and to conduct audits across systems. Worse still, they are inefficient. Many organizations have already made a significant investment over the past decade in identity governance processes, such as access request, access reviews and remediation of inappropriate access. If point solutions provide these capabilities on top of what already exists, the organization is making a redundant investment that deviates from the centralized approach.

The Right Way to Manage Unstructured Data for IAM Leaders

Given the severity of security and regulatory risk, unstructured data is an issue that cannot be ignored or addressed with legacy approaches. As an IAM leader for your organization, you must view the management of unstructured data as a key extension of your mission. Here are four practical steps and processes to make this shift happen.

Taking Control of Unstructured Data

Data is undoubtedly the focal point of unstructured data management, but in order to manage unstructured data effectively, identity governance expertise and processes are required.

Specifically, the IAM group plays a central role to ensure that:

- All high-risk, sensitive unstructured data has an appropriate access control model in place.
- Access to sensitive data is granted with oversight and according to pre-defined access policies.
- A periodic access review process by the appropriate business and technical owners is used to mitigate compliance and security risks.
- There is timely remediation of inappropriate access to unstructured data.

Build Bridges with Peer Data Management Groups

In order to effectively manage unstructured data, IT groups need to coordinate and collaborate across departmental boundaries. Instead of focusing on domain-specific plans, IT leaders must think in terms of a risk management strategy for the entire organization.

If you lead the IAM group, here are some practical tips to making the collaboration happen:

- Include the management of unstructured data in your identity governance mission statement, plans, and budget. Share this with your colleagues.
- Make sure executives understand identity governance is a key ingredient of securing unstructured data.
- If other groups feel they own the management of unstructured data, volunteer the identity governance group to collaborate with them on a complete solution.
- Be prepared to offer specific services the identity governance team can provide, including access reviews, access request, and remediation of inappropriate access.
- Share information about investment in people and tools. In the process, you'll be able to discover if redundant tools are being used as tactical "point" solutions.

Invest in Tools to Make Unstructured Data Manageable

Unstructured data presents some unique challenges for identity governance staff. It's important to know that you'll need to invest in complementary tools to analyze unstructured data in order to understand how it should be managed. To supplement the identity governance solutions you already have in place, here are the capabilities that you will need:

- **Sensitive data discovery and classification:** With huge volumes of unstructured data stored throughout the enterprise, it is important to define and prioritize risks to secure the most sensitive, business-critical data. The goal is to gain full visibility across all sensitive data, and ensure you have the right controls in place to protect it.
- **Data ownership and control:** Identifying the appropriate data owner is a critical success factor. Most organizations do not have an established framework for ownership of unstructured data, so you will need support for identifying potential data owners based on behavioral patterns or a data owner election process that allows for easy assignment of responsible data owners.
- **Entitlement analysis:** You will need support for analyzing the complex permissions that are used to grant users access to unstructured data, including nested groups in Active Directory and inherited rights in the file system.
- **Easy integration to existing identity governance systems:** Ensure that your tools for managing unstructured data can be quickly and easily integrated with your identity governance solutions so that you can leverage policies and processes already used for securing structured data and applications.

Say No to Management Silos

While you may be tempted to implement a tactical tool for managing unstructured data, you should not lose sight of the need to govern and control data access across your entire organization. By investing in a unified solution to manage access across applications, systems and data stored in files, you can ensure policy and process consistency, and you can avoid the wasted effort of duplicating identity governance functionality in domain-specific tools.

Even more importantly, to effectively reduce risk and to make your workplace secure, you need to see the big picture. You can't govern without centralized visibility to "who has access to what?" across both structured and unstructured data. And you cannot readily meet audit and compliance requirements without consistent access control policies and standards, access review processes, and user provisioning processes.

Take Action with SailPoint

As unstructured data continues to proliferate across enterprises at an unprecedented rate, IAM leaders can no longer ignore the inherent risks it poses to their organizations. Now is the time to tackle the problem. The good news is that with an identity governance strategy that spans applications, systems, and data, it's possible to mitigate these risks and chart a path towards more effective compliance and efficiency.

SailPoint identityIQ provides organizations with a comprehensive approach to securing access across all applications and unstructured data. Organizations can discover where sensitive unstructured data resides, establish access controls, and gain real-time visibility across on-premises and cloud storage systems. With a holistic identity governance program, IAM leaders can better address security threats, more confidently ensure compliance, and empower the business by making sure the right people have the right access to the right information.

SAILPOINT: THE POWER OF IDENTITY™

sailpoint.com

SailPoint, the leader in enterprise identity management, brings the Power of Identity to customers around the world. SailPoint's open identity platform gives organizations the power to enter new markets, scale their workforces, embrace new technologies, innovate faster and compete on a global basis. As both an industry pioneer and market leader in identity governance, SailPoint delivers security, operational efficiency and compliance to enterprises with complex IT environments. SailPoint's customers are among the world's largest companies.

© 2019 SailPoint Technologies, Inc. All rights reserved. SailPoint, the SailPoint logo and all techniques are trademarks or registered trademarks of SailPoint Technologies, Inc. in the U.S. and/or other countries. All other products or services are trademarks of their respective companies.

WP1191-1903