AvePoint

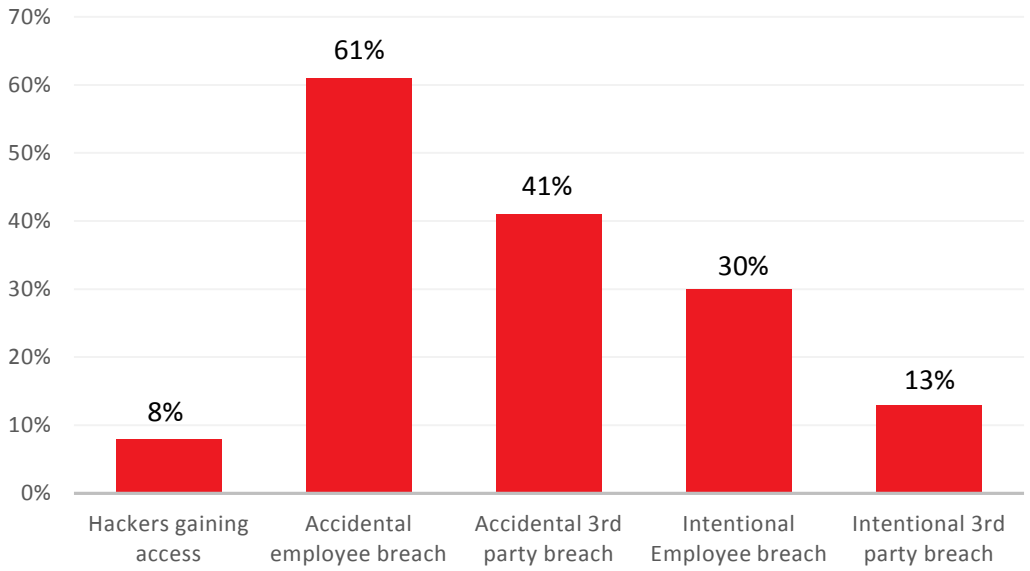
PREVENTING THE NEXT GLOBAL PRIVACY BREACH: BEST PRACTICES FOR DATA DISCOVERY

TABLE OF CONTENTS

INTRODUCTION

There hasn't been a week that's gone by recently where a company avoids the spotlight for losing individuals' sensitive information. From Edward Snowden to Target and now Anthem Healthcare, these data and privacy breaches are tremendously damaging – the ramifications of which haven't fully come to bear yet. What these high profile cases don't shine a light on, however, can be just as damaging: Company employees continue to cause more data breaches, on average, than anyone else. Even worse? In many cases, it's purely accidental.

To illustrate this concern, as reported in a survey of 518 responses[1] from a variety of public, private and not-for-profit companies, the following key findings were found:



Why is accidental breach so common? There is a vast dichotomy between business realities and privacy needs - today. Everyone is a content contributor – being asked to collaborate more, produce faster, and innovate in order to drive business initiatives. We're removing speed limits and mandating people do what they need to in order to get their jobs done.

While we're empowering information workers to use data quickly to perform tasks, they're not armed with the right information or training to do so safely. Who is ultimately responsible? If your organization has less than 20,000 employees, your security team is on the hook for addressing these concerns. Larger companies may also utilize Chief Privacy Officers and Compliance Officers, but a recent report from Forrester Research[2] found that

[1] http://www.bentley.edu/centers/sites/www.bentley.edu.centers/files/centers/cbe/cbe-external-surveys/data-privacy.pdf
[2] Understand the State of Data Security and Privacy: 2013 to 2014, Heidi Shey, Forrester Research

security teams are “fully responsible” in 30 percent of organizations for privacy and regulations, and 34 percent of enterprises hold the security team as “mostly responsible”.

This is a tremendous burden, but also a tremendous opportunity for security and privacy teams to help enable the rest of the organization to collaborate, contribute, and innovate in ways that are safe for not only the organization but that of the customers, partners, and external vendors who provide organizations with sensitive information and expect it to be managed properly. The overarching goal is to ensure information is available to those who should have access to it, but protected from those who shouldn’t.

Best practices for security and risk management have traditionally focused on “building walls” around the perimeter to “keep people out” and “keep information in”. However, the challenge with this approach is that as you build a ten foot wall, your opponent brings an eleven foot ladder. Thus you are always in a defensive mode, looking to outwit an enemy.

There are many methods of assessing risk – ranging from a flip of a coin to a much more prescriptive, mathematical approach. Perhaps the most important thing to consider is what is actually considered risk. Analysis of this risk required a balance of standards, exposure, and what that means to your business. Organizations that have a better handle on risk analysis factor this risk identification into classification schemes used to file data. In a simple example, content classified with a higher risk is separated and managed differently than content at a lower risk.

A robust risk management program not only involves surfacing or identifying risk, but should also include the ability to audit and limit the risk. It’s imperative to rate the risk and the likelihood of being impacted by the same as well as the real impact of one risk weighed against others.

Generally speaking, organizations should look to use technologies and create policies that make information available to the people that should have it and protect it from the people who should not. With highly sensitive data (Personally identifiable information, Protected Health information, and others), limited and appropriate access is always critically important. Simply put, understanding the difference between what can be shared and what should be shared is always the key.

In this white paper, we’ll focus on what is arguably the most important step in your risk management program: data discovery. It’s impossible to fix problems you don’t know exist. Many enterprises have file shares that are home to terabytes or petabytes of information, with no real organization or ability to audit it. We’ll explain the case for file analysis tools, its usage, and best practices for data discovery and defensible destruction.

WHERE DO YOU START? FILE ANALYSIS.

While perimeter-based security is important, it is only one strategy in an approach that must be layered. Organizations must also look at information as it is managed throughout their information gateways. At rest or in motion, data flows through file shares as well as Web sites, Web applications, SharePoint sites, communication systems (such as email and instant messaging) and social media platforms. By thinking

holistically about managing compliance as well as maintaining visibility, data classification, and control as information moves about the organization, the data walls become less and less penetrable.

In the Gartner Research's Market Guide for File Analysis Software[3], enterprises realize they need to understand their data better in order to not just facilitate better use of it, but to manage growing storage environments. The market guide outlines the three primary reasons file analysis tools are implemented:

1. Increase operational efficiency
2. Lower costs
3. Mitigate corporate risk

The theory here is by identifying and classifying the unstructured data (think Word documents, PDFs, videos, images, and the like), organizations can make more informed decisions regarding risk identification. You can also get a better handle on identifying which documents to keep and which can be archived and/or destroyed, which reduces storage costs, document clutter, and eases a transition to newer document sharing platforms if this is part of your roadmap. Back to the privacy concern, file analysis tools can reduce the risk of privacy breaches because you have a much better understanding of where the files are stored and who has access to them.

By investing in compliance-centric classification strategy, companies can turn what was previously considered to be a "cost to the business" (security, privacy, and DLP technologies) into a corporate asset. It's very important to understand just how much risk you have in your organization today. Knowing you could be on the hook for millions of dollars in fines is important to understand at every level of the organization. But in mitigating this risk through a compliance solution, organizations can also find where the dark data lives across your enterprise. Reinforcing this point, this also drives enterprise classification and taxonomy across systems.

We've empowered business users with "spell checking" in most authoring systems, and we expect that communications sent for business purposes will not have spelling mistakes. By empowering organizations and particularly decision makers with information about where sensitive data lives – and how the business is truly managing data flow across the organization – decisions can be made quicker and more efficiently to enable rather than block productivity.

Not knowing is never better – so understand how your business users are managing sensitive data today so that you can properly remediate and educate going forward after a system clean-up.

For effective data management and collaboration to turn into a competitive advantage for the business, timely access to data as well as multi-directional communication flow – with the right risk management filters in place – is essential so that data is available whenever and wherever to those who need it, and not available to those who shouldn't have access. Companies can repurpose their compliance programs that have traditionally been viewed as a "cost center" for the business, to help them turn this previously untapped information into a business asset. This not only creates a quantifiable return on investment for data security and privacy programs, but also helps the company increase productivity and mitigate regulatory compliance issues.

[3] Market Guide for File Analysis Software; Alan Dayley, Garth Landers, Debra Logan, Earl Perkins, Gartner Research

MOVING TO THE CLOUD

Organizations worldwide are quickly adopting cloud computing in order to stay ahead of a hyper-competitive business landscape. The immediate draw to cloud computing is clear – reduced total cost of ownership and less hardware for IT administrators to maintain. Hosting your documents on cloud platforms including Box or Microsoft Office 365 may reduce cost and improve global access to content. However, for organizations subject to regulatory requirements, the move to the cloud is not without risk. Enterprises have significant concerns about storing business data outside the walls of their enterprises, due to a number of reasons:

- Perceived risk that data is more vulnerable, increasing risk of inappropriate access
- Non-employee IT administrators possessing a high level of access and control over information
- Technology choices and complexity associated with authentication and authorization
- Regulations around data sovereignty

To mitigate these concerns, organizations can opt to offload select content or certain workloads to the cloud, and keep their most regulated content on-premises.

Content previously stored on file shares can now be revitalized and extended to the enterprise stored in a new repository on the cloud. Before moving that data, ensure that your in-house organizational policies and practices are properly extended to your cloud-hosted applications.

As companies and government agencies move their applications increasingly to a cloud-based infrastructure, they must also understand and fully review the associated privacy and security considerations. Privacy is a global issue, and one thing is certain: You cannot ignore privacy, as your internal and external customers will demand it.

BEST PRACTICES FOR USING FILE ANALYSIS TOOLS

Data-aware security policies provide an opportunity for organizations to build a more layered approach to security, prioritizing where effort (and costs) should be spent, and building multiple lines of defense to create a defensible strategy for data destruction.

While there are many information gateways an enterprise uses at any given time, this white paper is focusing on file shares. Found in virtually every organization, most of the data could be quite old, and in fact potentially left there by employees who are no longer with the company (and pre-dated current compliance and security policies). A substantial amount of it may be deemed sensitive and/or confidential, so unauthorized access and/or disclosure of it could expose an organization to legal liability. Many organizations also look the other way when it comes to the e-discovery hazard with their file shares. Noting that the average cost of performing an e-discovery exercise can average $18,000 USD per gigabyte of content[4], allowing these digital landfills to continue to grow only increases this effort and cost.

[4] http://www.insidecounsel.com/2012/05/23/e-discovery-costs-pay-now-or-pay-later

As we previously mentioned in this white paper, a file analysis tool helps discover data, classify it according to your specific business needs, and then enables you to review and take appropriate actions on the data – whether it means moving reclassifying it, it, or deleting the information. Broken down into four discrete steps, here are our best practices when using file analysis solutions.

STEP 1: CLASSIFY YOUR DATA

The first step to take in properly disposing of data is to determine what type of data you have. An example of a common data classification schema is data must be classified as public, internal, sensitive, or restricted. The classification of the data dictates its disposal method. This does not have to be completed as an all-or-nothing effort, but rather can be done through a phased approach and as part of an initial "discovery" project across a limited scope of data to help build the business rules that can then be disseminated across the organization's data repositories. The goal of the classification in this step is to get far enough along so that you can proceed with the second step.

STEP 2: DETERMINE RETENTION SCHEDULE

Once you have determined the classification of your data, you need to be sure that it is not subject to any retention periods. Region and government-specific laws and regulations, requirements of accrediting and other external agencies, and prudent management practices govern the retention and disposal of organizational records. These records must be retained appropriately and disposed of in a timely manner to meet the requirements of external regulations.

STEP 3: ASSIGN HISTORICAL VALUE

After you have determined that data in your possession is not subject to any retention period, it is important to evaluate whether the documents have any historical or archival purpose for the organization. In some instances data ready to be disposed may contain information with enduring legal, fiscal, research, or historical value, and should be retained and preserved indefinitely.

STEP 4: DISCARD FILES

Finally, after data in your possession is classified and reviewed for retention and archival purposes – and it is determined that the data can be properly discarded – the last step is to dispose of your data in the appropriate manner.

CONCLUSION

With the increased frequency and damaging data loss due to privacy breaches, what can you do to avoid being the next statistic? For starters, remaining passive is not a choice and not knowing is never an acceptable excuse. Don't fear the dark: Take action to shine a light on the dark data living in your file shares. By discovering, mapping, and classifying the unstructured files, organizations can make more informed decisions regarding

which data to keep and remove. This also eases a transition to newer file management systems that are more suited for tagging and managing content throughout its lifecycle.

A file analysis exercise is able to reduce the risk of privacy or sensitive information breaches because you can identify where the files are and who has access to them. As organizations spend big money on big data, they should look inwards into the data they already own. This solution also creates the opportunity to realize and take advantage of the full potential of the "big data" that is stored in vast existing repositories.

Bearing this situation in mind, AvePoint's File Analysis solution provides the ability to completely integrate data scanning, tagging, record retention, archiving, and data loss prevention into your daily workflow. By empowering organizations – and particularly decision makers – with information about where sensitive data lives and how the business is truly managing data flow across, decisions can be made quicker and more efficiently to enable rather than block productivity. To learn more, visit www.avepoint.com.