Osterman Research

WHITE PAPER

White Paper by Osterman Research
Published **February 2019**
Sponsored by **OpenText**

Best Practices for File Governance

Executive Summary

"File governance" can be defined as the management of all files in an organization, including how they are protected from data breaches and other security threats, how and where they are stored, how they are accessed, how they are otherwise managed, and how they are deleted when no longer necessary. File governance includes all of the policies, processes and technologies that are involved in ensuring that these files are properly managed according to corporate, regulatory and legal requirements.

KEY TAKEAWAYS

- Although there are a number of important drivers for good file governance, compliance is among the most important drivers. Regulations like the European Union's (EU) General Data Protection Regulation (GDPR), have been the catalyst that is motivating many organizations to get their file governance house in order.
- Security is another critical driver for good file governance, since it can help organizations to recover rapidly from ransomware infections, prevent data breaches, and prevent access to sensitive and confidential information by unauthorized parties.
- Proper file governance is an essential element of any organization's security posture, since good file management can significantly reduce the potential for data breaches and can mitigate the impact of ransomware attacks.
- While many information managers are much more willing to retain files than delete them, defensible deletion is an essential element of any good file governance program and can reduce an organization's exposure to risk by reducing the amount of data that is subject to breach.
- Good file governance requires a change of mindset, including viewing data as both a critical asset and a major security risk.
- Automation of file management activities, particularly for larger organizations, is an essential element of any good file governance program.
- There are a number of best practices that organizations must adopt as part of their file governance evaluation, including understanding where their files are located, analyzing the content of files, understanding the risk posture of the organization, and defensibly deleting unnecessary content.

Good file governance requires a change of mindset, including viewing data as both a critical asset and a major security risk.

ABOUT THIS WHITE PAPER

This white paper was sponsored by OpenText; information about the company is located at the end of this document.

What Does Good File Governance Really Mean?

COMPLIANCE IS BECOMING INCREASINGLY CRITICAL

"Compliance" is a broad term that covers a wide range of file governance issues across a range of industries and geographies. While heavily regulated industries, such as financial services and healthcare, must comply with a variety of fairly strict compliance regulations, compliance has become an issue for virtually every type of organization, regardless of the industry it serves. For example:

- Various types of organizations in the financial services industry must retain data for multiple years, and so must retain a variety of different data types – email,

social media posts, text messages, advertising content, and so forth – for multiple years. In the United States, for example, broker-dealers, investment advisers, hedge fund managers and others must retain data for several years, more recent data in systems that make the data easily and quickly accessible. Penalties for non-compliance with Securities and Exchange Commission (SEC) and Financial Industry Regulatory Authority (FINRA) requirements can be significant.

- The Financial Conduct Authority (FCA) in the United Kingdom imposes various requirements on the tens of thousands of financial institutions in the UK. These include the Senior Management Arrangements, Systems and Controls (SYSC) rules (SYSC 3.2.6R and 6.1.1R), requiring the deployment and maintenance of systems designed to detect and mitigate risks associated with financial crimes; and Principles 2 and 3 of the Principles for Businesses that requires the application of appropriate risk management capabilities that are commensurate with the risk of financial crimes that might be perpetrated against customer data.

- The Payment Card Industry Data Security Standard (PCI DSS) is a set of requirements for protecting the security of consumers' payment account information, such as credit card data. It includes provisions for encrypting cardholder data when it is sent over public networks, building and maintaining a secure network, and assigning unique IDs to each individual who has access to cardholder information.

- Healthcare institutions in the United States must adhere to the Health Insurance Portability and Accountability Act (HIPAA) mandates, which require protection of individually identifiable health information, or Protected Health Information (PHI). These apply to both "covered entities" – those providing direct care – and "business associates", of which there are many and varied types, such as benefits administrators and insurance companies. The HIPAA Security Rule requires that healthcare institutions implement appropriate administrative, physical, and technical safeguards to assure the confidentiality, integrity, and availability of PHI, which is most often held in individual files. For example, if data has to be sent to another person or institution and there is a significant risk of unauthorized disclosure, data encryption will be required.

- The energy industry – e.g., electrical service providers, natural gas providers, oil products providers, energy traders, public utilities and non-utility energy producers – are subject to a variety of strict compliance obligations. For example, Federal Energy Regulatory Commission (FERC) Rule No. 717 requires that all communications between marketing-related and transmission-related employees of a vertically integrated provider be retained for long periods of time. Other FERC requirements focus on retention of data like minutes of stockholder-related meetings, plant ledgers and other types of data.

Aside from the large number of industry-specific compliance obligations, of which the above examples are only a small representation, there is a growing variety of consumer privacy-focused regulations that represent compliance for organizations that previously would not have considered themselves "heavily regulated". For example:

- The GDPR sets out a variety of compliance obligations for organizations that process or control data on residents of the EU. These obligations include compliance obligations for how data is collected, how long it is retained, the types of data that may or may not be retained, how data is shared among processors, and so forth. Fines for non-compliance are, by some measures, draconian and could result in penalties of several billion dollars in some cases.

- Several other types of compliance obligations, similar in scope to the GDPR, are being proposed or coming into force in the near future. These include the California Consumer Privacy Act, India's Personal Data Protection Bill of 2018, the

There is a growing variety of consumer privacy-focused regulations that represent compliance for organizations that previously would not have considered themselves "heavily regulated".

new General Data Privacy Law 2018 in Brazil, and Australia's new data breach notification law.

- Every US state now has a data breach notification law that requires notification to residents of that state if certain types of their data are lost or stolen.

LEGAL CONSIDERATIONS ARE ALSO A FORM OF COMPLIANCE

There are a number of legal considerations for the proper governance of files and other data that constitute a set of de facto compliance obligations for virtually any organization, albeit with sometimes less specific requirements and less definite retention periods than what are generally imposed by codified compliance regulations. For example:

- Organizations in the United States are obligated to retain relevant data if they anticipate that they might be party to a legal action, even if no formal litigation has begun. For example, if senior managers reasonably anticipate that their firm will be sued by a terminated employee, it is incumbent upon the organization to retain all documents that might be relevant in that legal action. This generally includes relevant emails and files, but can also include text messages, social media posts and other content.

- Organizations that are presented with an eDiscovery order must search for, find and produce court-ordered content in a timely way. This means that emails and files are very likely to be the subjects of a discovery order, but various other types of electronic content may also be required.

GOOD FILE MANAGEMENT IS AN ESSENTIAL ELEMENT OF GOOD SECURITY

Preventing ransomware is a leading priority for IT decision makers in most organizations given the debilitating impact that ransomware can have. There have been thousands of major ransomware attacks worldwide over the past few years that have shut down or seriously impacted hospitals, shipping companies, police departments, city governments and other organizations.

In an ideal world, organizations would be able to prevent ransomware from infecting their endpoints and networks. However, given that even the best security solutions and security awareness training will not guarantee complete blockage of ransomware exploits, it is essential that organizations have in place robust archiving and backup technologies that will enable rapid recovery from a ransomware infection to a known good, pre-infection state. This is essential to avoid the myriad costs that go into recovering data and reimaging infected machines, not to mention the potentially enormous hit on employee productivity and the ability to transact business. Good file governance can go a long way toward minimizing the impact of ransomware in any organization.

It is essential that organizations have in place robust archiving and backup technologies that will enable rapid recovery from a ransomware infection.

FILE GOVERNANCE AND EMPLOYEE PRODUCTIVITY GO HAND-IN-HAND

It's important to note that proper governance of an organization's files is an enabler for employee productivity. For example, if files are in IT-managed locations, IT understands where all of the organization's files are located, and IT provides robust search capability for this content – in short, they have a good file governance program in place – then employees will spend significantly less time searching for information than if good file governance is not the norm. The average employee spends 9.3 hours per week searching for information[1] – if good file governance has been implemented, employees will spend significantly more time on productive work.

[1] http://www.bwd-it.com/various-survey-statistics-workers-spend-too-much-time-searching-for-information/

DEFENSIBLE DELETION IS ESSENTIAL AS PART OF INFORMATION GOVERNANCE

The decision to retain files is considered a "safe" decision because this content will be available for compliance, legal and other reasons. The decision to delete files – while not inherently unsafe – is often avoided. However, most organizations have enormous volumes of old files that are no longer necessary or are duplicates of other files and so can safely be deleted, thereby avoid the cost of storing this content. Moreover, eliminating files can substantially reduce risk, since there is less data that can be lost if a data breach occurs.

ACCESS CONTROL IS ESSENTIAL FOR GOOD STEWARDSHIP OF FILES

A key element of good file governance is maintaining appropriate controls on who can access data. For example, every employee will rarely, if ever, need to have access to every file in the organization, and the appropriate controls should be in place that will enable access to files to only those who need it. Moreover, any file access solution should include the ability to monitor file access so that anomalous activity – such as an employee who is accessing large volumes of data all of a sudden – can be investigated quickly and easily.

Some Important Issues to Consider

VIEW DATA AS AN ASSET…AND A SECURITY RISK

Data is a critical asset and, for the majority of organizations, their single most important one. Moreover, the largest single source of content is in files, whether stored on-premises in file servers or in the cloud. Consequently, files should be treated as the critical resource they are on a number of levels:

- Robust security should be implemented to protect files from the variety of threats that are directed toward them, such as malware that could result in a breach of this data and phishing attempts that could deliver that malware.
- Files that contain even slightly sensitive or confidential data should be encrypted in transit and at rest.
- Access management should be implemented so that only those who need to see particular files should have access to them.
- Monitoring should be implemented to ensure that unusual file access or disposition is quickly detected and addressed.
- File analytics should be implemented to ensure that decision makers understand the content within files – those that contain data that are subject to various compliance obligations, that contain data that might be in violation of corporate policy, that need to be retained and currently are not, that must be encrypted, that contain obsolete data and can be safely deleted, or that are rarely accessed and can be moved to less expensive storage tiers.

A key element of good file governance is maintaining appropriate controls on who can access data.

A SCALABLE ARCHIVING CAPABILITY IS A MUST

Any file governance program must include at its core a robust and scalable archiving capability to ensure that content can be archived for as long as necessary, and so that the files can be searched and produced quickly and easily. It's essential that any archiving capability cover not only files, but email and any other content that the organization must or should be archiving, such as social media posts, text messages and the like.

A key element of a good archiving solution will be the ability to move files to the appropriate storage tiers. For example, files that are frequently accessed should be stored on a higher level storage tier that provides quick access, whereas files that are rarely accessed should be stored on a lower level storage tier that is less expensive. This is particularly true for cloud-based storage, since providers like Amazon Web Services charge substantially less for low-level storage (e.g., Glacier, priced at $0.004 per gigabyte) than they do for high availability storage (e.g., S3 Standard, priced at $0.021-0.023 per gigabyte)[2].

FIND HOLES IN SECURITY TO AVOID DATA BREACHES

It's essential that to understand how files are managed in an organization – and make appropriate changes – to ensure that the opportunity for a data breach is minimized. Important questions for IT and business decision makers to ask themselves include:

- Are files stored in venues that are both IT approved and accessible to IT in the event all files need to be accessed, such as for a Subject Access Request under the GDPR?
- Are policies in place to ensure that employees know where and how they should store files instead of storing them wherever it's most convenient?
- Are some files accessible via the Internet when it might be smarter to store them off-line and make them completely invulnerable to theft?
- Are files that are accessible on-line encrypted if they contain sensitive or confidential information?
- Are files accessible only to employees who need them, or can any employee access any file on the network?

BE ABLE TO SURVIVE AN AUDIT

The "acid test" for any good file governance program is the ability to survive an audit, whether it's from a regulator like FINRA, an internal audit from the compliance department, or a quasi-audit delivered via a court order asking to produce certain types of information during a legal action. If an organization is able to produce what it needs to produce in the timeframe allotted, that will indicate that it can satisfy future compliance requirements with minimal risk of failure.

The "acid test" for any good file governance program is the ability to survive an audit.

AUTOMATION IS KEY IN A LARGE ORGANIZATION

Automation is an essential element for organizations that manage large numbers of files. For example, if we assume that a 500-user firm maintains storage of one terabyte of data, and that there are 3,000 files per gigabyte, that means that the organization will store roughly three million files. Managing even this many files – not to mention the tens or hundreds of millions of files that a much larger organization would maintain – is simply not tenable without the use of automation.

Human management and classification of content, while useful, is not always reliable and can result in significant mistakes. For example, if a company hires a new employee in Detroit, it might classify his or her application and health insurance data in a particular way to satisfy US compliance requirements. But if that employee is transferred to Paris three years later, GDPR requirements will now take effect where they did not before, and the appropriate files will need to be managed differently. Without automation, conventional file-management processes are unlikely to make the appropriate changes and the organization will be out of compliance. In short, file monitoring using automated processing is an essential element of any good file governance program.

[2] Based on AWS pricing in an Ohio-based data center (https://aws.amazon.com/s3/pricing/?nc=nsb&pg=sc)

FILE GOVERNANCE INVOLVES BUSINESS MANAGERS, NOT JUST IT

Finally, it's important to remember that file governance is not so much an IT issue as it is a business issue. While IT is typically charged with implementing the archiving solution and all of the other elements that are part of a good file governance solution, the need for good file governance should be driven by business decision makers with input from all stakeholders, including IT.

Next Steps

The appropriate best practices for an organization will depend on a number of factors, including the industries they serve, the regulations with which they must comply, the legal environment in which they operate, the risk tolerance of their senior decision makers, and the technologies and processes they have or have not implemented. However, Osterman Research offers the following recommendations for any organization to consider as it develops its file governance program.

CONDUCT A FILE INVENTORY

Most organizations would be at a loss if asked if they could identify the location of every file within their organization. Because files are stored on file shares, in various cloud applications, in SharePoint, on employees' PCs, on mobile devices, on employees' home computers and often in many other locations, it is essential to determine where all of an organization's files are located. Doing so will enable an organization to understand where it has duplicate files, where it has files in unprotected locations that contain sensitive data, and how it can get a handle on controlling the sprawl of files across the organization.

Most organizations would be at a loss if asked if they could identify the location of every file within their organization.

ANALYZE YOUR CURRENT FILES

An important next step is to use advanced analytics to analyze the files currently within the organization to understand exactly what they contain. It is very likely that many of these files contain sensitive information like customer payment records, employees' personal information, and other sensitive and confidential information that must be encrypted or otherwise protected. Performing analytics on this data can also enable decision makers to understand which files should be brought into compliance with the organization's records retention schedule, but that might not be.

UNDERSTAND YOUR RISKS

The file inventory and analysis will help decision makers to understand what files are currently posing a risk to the organization because they are not encrypted, not classified properly under the appropriate retention schedule, or simply not necessary to retain any longer.

DEVELOP OR ENHANCE RETENTION SCHEDULES

Most organizations have some sort of records retention schedule, but the analysis might yield that current retention policies are not sufficiently granular or have not been updated to comply with new compliance obligations. This would be a good opportunity to develop new retention schedules or at least enhance and improve those that already exist.

DEFENSIBLY DELETE WHAT YOU CAN

Most decision makers are much more willing to retain data for long periods than they are to delete older content that is no longer necessary. However, easily 20 percent of the files that an organization possesses can safely be deleted, and in some cases it might be 50 percent or more. If a file no longer serves a useful purpose to the organization, is not required to be retained, or has not been accessed for an extended period, it can very likely be safely deleted as long as the deletion process is well documented.

IMPLEMENT THE RIGHT TECHNOLOGIES TO MANAGE FILES MOVING FORWARD

After the above steps have been completed, it's essential to deploy the right technologies that will enable proper file governance moving forward. Of course, these include appropriate archiving technologies that will enable archiving of files and other content, but also monitoring technologies that will detect when files change state (e.g., from one that does not need to be managed according to compliance rules to one that does), access management, encryption solutions, etc. In short, file governance is not a project, but an ongoing process.

Summary and Conclusions

Good file governance is a necessity if an organization is to protect its most valuable data assets from security risks like ransomware and data breaches, and from non-compliance with its regulatory and legal obligations. Consequently, every organization must evaluate its current position with regard to the risks it faces from inadequate file governance and take the appropriate steps to ensure that it has reduced its risk to the greatest extent possible, while also ensuring that it maximizes employee productivity and control over its data assets.

About OpenText

OpenText™ is The Information Company. We power and protect information to elevate every person and every organization to gain the information advantage. A leader in global Information Management, OpenText offers a comprehensive portfolio of solutions across content, business network, digital experience, security, application modernization, operations management and developer APIs. OpenText solutions help customers simplify their systems, connect their data, build frictionless automation and thrive in a multi-cloud world. The company fosters inclusive environments that leverage the diverse backgrounds and perspectives of all employees, customers, suppliers and partners. For more information about OpenText (NASDAQ/TSX: OTEX), visit www.opentext.com.

OpenText™ Portfolio solutions help organizations know their data, empower their people, and drive their future. Automated compliance solutions provide real-time data analytics and privacy reports. Productive, empowered people achieve flexible, smarter, more collaborative work environments. Give remote workers the right content, for the right people, at the right time, on any device. Learn more at www.opentext.com/products/digital-workplace.

It's essential to deploy the right technologies that will enable proper file governance.