Best Practices for the Next Generation of Content Archiving and Content Services

An Osterman Research White Paper
Published April 2018

Viewpoint*e*

Executive Summary

- Decision makers need to understand that the traditional role of archiving is defensive. Early generation archiving solutions are simply not designed for extracting business intelligence and insight in a meaningful way.

- The traditional, defensive roles for eDiscovery and compliance will continue indefinitely, but new, proactive applications that extract insight and intelligence from content are essential for any information-focused organization.

- For organizations that are archiving email today, decision makers need to consider how to retain and manage content outside of traditional archives and how to extract insight and intelligence from the content.

- Key issues to address for next-generation archiving or content services solutions include risk mitigation and gaining competitive advantage from the proper management of retained information.

- There are a number of data types that organizations consider to be important to manage and use on a long-term basis. The five most important data types that are currently archived are:

  o Corporate email
  o Content from departmental file shares
  o User-generated files
  o Accounts payable content
  o CRM data

- eDiscovery and compliance are the leading drivers for archiving and content management because of the frequency of lawsuits and compliance audits. Enterprises need a way to preserve all types of electronic content and they need to be able to search it quickly.

- The cloud is a key destination for enterprise applications and it directly impacts content management. We found that today 32 percent of total enterprise content is stored in the cloud, but in two years 52 percent of it will be stored in the cloud – a 62 percent increase in the space of just 24 months.

- New, next generation archiving or content services solutions have two major differences compared to traditional email archiving solutions:

  o First, they support a wide range of content types, not just email.

  o Second, they support content that is stored on premises <u>and</u> in the cloud.

- There are a number of important use cases for using electronic content for business intelligence purposes, among which are:

  o Improving the sales and support processes
  o Bringing employees up to speed more quickly
  o Preserving content when employees leave
  o Conducting investigations
  o Enabling employees to find others more easily

The cloud is a key destination for enterprise applications and it directly impacts content management.

The Importance of Different Data Types for Long-term Retention

The key to better content management is to understand the importance of each content type to the organization. In the survey conducted for this white paper, we queried 157 decision makers and influencers in mid-sized and large organizations, primarily in North America. The goal of the survey was to understand the current and future role of electronic content management and archiving, and their impact to the organization.

One of the primary issues that we wanted to address was the importance of various data types that organizations use and manage on a long-term basis. As shown in Figure 1, the most important data types in this regard are corporate email, content from departmental file shares, user-generated files, accounts payable content and CRM data.

Figure 1
Importance of Managing and Using Data Types on a Long-Term Basis
Percentage Responding Important or Extremely Important

Content Type	Low Import-ance	Moderate Import-ance	High Import-ance
Corporate email	1%	20%	79%
Content from department file shares	3%	24%	73%
Users' files	4%	25%	70%
Accounts payable content	2%	28%	70%
CRM data	7%	31%	62%
Content from SharePoint or similar collaboration tools	9%	39%	52%
Content from company-managed file sync and share tools, e.g., Dropbox	7%	43%	50%
SAP data	27%	31%	42%
Content from image capture systems	17%	43%	40%
Content from company surveillance video	19%	42%	38%
Content from company-owned mobile devices	16%	46%	37%
Voicemails from the company phone system	26%	43%	31%
Content from ex-employees	20%	50%	30%
Corporate instant messaging/mobile message app content	26%	48%	26%
Work-related content from employees' personal mobile devices	31%	46%	24%
Content from users' personally managed file sync and share tools	32%	45%	23%
Corporate LinkedIn posts	32%	47%	21%
Corporate Facebook posts	33%	46%	20%
Corporate Twitter content	34%	48%	18%
Text messages	36%	46%	18%

Content Type	Low Import-ance	Moderate Import-ance	High Import-ance
Work-related content from employees' instant messaging accounts	39%	45%	16%
Work-related posts from employees' personal social media accounts	38%	46%	16%
Other	19%	55%	27%

Source: Osterman Research, Inc.

The list of different content types reveals the wide number of new, emerging content types that organizations consider to be important. Examples are content from company managed file sync and share tools, image capture systems and company surveillance video. It is interesting to note the ranking for the popular social media content types – LinkedIn, Facebook and Twitter. As compared to other content types, these are generally of least importance to decision makers.

Email Archiving

First generation email archiving solutions were designed primarily to manage mailbox size. In the early days of email, mailbox size was limited to only tens or hundreds of megabytes. These email archiving solutions were designed to remove large email (and attachments) and replace them with a small "pointer" to the archive. This feature allowed users to keep months of email without exceeding the mailbox size limit. Today's modern email solutions, such as Office 365, support multi-gigabyte mailboxes, capable of holding significantly more email data than earlier solutions. As a result, mailbox size management is less of a driver for email archiving than it used to be and we see its importance in the context of archiving continuing to dwindle. That doesn't mean that users no longer need archiving for mailbox management. Some users continue to run into mailbox size limits – this is especially true for users who employ email as their primary file-sharing solution.

In the early 2000s, legal discovery and regulatory compliance emerged as a new driver for email archiving. Because email can be deleted by users, enterprises needed a way to preserve email in a dedicated repository where it cannot be modified or destroyed. As email archive solutions developed, new content types were added, including instant messaging, files, SharePoint and social media, increasing the number of content types that can be searched for legal discovery. All archived electronic content is stored in a dedicated archive repository that traditionally has been managed on premises. However, a growing proportion of archive solutions – and a growing share of corporate data – support cloud storage.

For the most rigorous preservation of email content, journaling is used. As its name implies, journaling keeps a copy of all email sent and received for each mailbox. It is the responsibility of the archiving solution to protect the journal email copy. In Office 365 environments, for example, an Exchange Online mailbox cannot be designated as a journaling mailbox. However, for organizations running an Exchange hybrid deployment with mailboxes split between on-premises servers and Office 365, administrators can designate an on-premises mailbox as the journaling mailbox for Exchange Online and on-premises mailboxes.

First generation email archiving solutions were designed primarily to manage mailbox size.

CURRENT ARCHIVING ADOPTION

To establish a baseline regarding current archiving and content management practices, we wanted to know about the types content that are archived today. As shown in Figure 2, corporate email, users' files and CRM data top the list of content that organizations archive.

Figure 2
Types of Content Archived Today

Content Type	%
Corporate email	93%
Users' files	80%
CRM data	63%
Content from departmental file shares, e.g., Legal, Finance	52%
Accounts payable content	48%
Content from SharePoint or similar collaboration tools	48%
SAP data	43%
Content from company-managed file sync and share tools, e.g., Dropbox or OneDrive	39%
Content from ex-employees	36%
Content from company-owned mobile devices	30%
Voicemails from the company phone system	27%
Content from company surveillance video	25%
Content from image capture systems	21%
Corporate Twitter content	20%
Text messages	20%
Corporate Facebook posts	18%
Corporate instant messaging/mobile message app content	18%
Work-related content from employees' personal mobile devices	14%
Corporate LinkedIn posts	13%
Work-related posts from employees' personal social media accounts	11%
Work-related content from employees' instant messaging accounts	7%
Content from users' personally managed file sync and share tools	5%
Other	13%

Source: Osterman Research, Inc.

It is no surprise that email is ranked as the leading type of content that organizations archive. Reasons for this include:

- Email is extremely valuable to employees for personal productivity and to the enterprise as a record of its business dealings. Osterman Research surveys of corporate end users have found consistently that information workers employ some aspect of email for about 150 minutes during a typical workday. As a result, an enormous amount of corporate information is bound up in the typical email system.

- Second, email has a high "velocity", meaning it is created quickly, moves quickly and is easily deleted. As a result, email archiving is a necessity to preserve email for eDiscovery and regulatory compliance. Modern enterprise email solutions support built-in archiving and, optionally, email journaling.

After email, user files rank as the second most common content type for archival. Files are managed on desktops, laptops and file shares, and an increasing number of files are being managed in cloud-based, file sync and share applications. Access to files is managed with passwords and files stored on file shares are backed up nightly, giving organizations the choice of archiving files or managing them "in-place".

Application data is managed by the source application and is also protected with passwords and backed up nightly. Generally, files and application data can be managed "in-place" without having to move or copy content to a dedicated archive, and so many decision makers are opting to maintain these critical data sources in their native location or application.

Data from image capture systems and company surveillance video may require special retention policies for company security reasons. Video files are quite large and create a storage challenge for long-term retention, and are difficult to move from on-premises storage to cloud storage, if so desired.

Archival of corporate social media and employee work-related social media and instant messaging accounts rank among the least archived data in mid-sized and large organizations. There are a couple of explanations for this:

- First, social media content is managed by the social applications themselves. Many organizations rely on the fact that content is protected and is accessed using the account credentials inherent in the platform when required for eDiscovery or regulatory compliance. However, there is no guarantee that social media content will be retained adequately by these providers, and so organizations that need to retain social media data, such as financial services firms, should have their own social media archiving capability.

- Second, a special connection (available from third-party archiving vendors) is required to archive content from social media accounts. This option adds cost and complexity and there may not be sufficient demand to warrant the IT investment aside from those industries in which retention is mandated.

Osterman Research takes the position that any electronic source of information – including social media, text messages and other data types that traditionally have not been archived – should be archived if it contains business information, or information that might be needed at a future date. For example, many business conversations include email, file transfer from a file-sharing platform, conversations via social media, and text messages, all of which should be archived if they contain relevant business content. Employees' social media, if used on a corporate network, should also be archived if these accounts are used by employees for business purposes.

The Drivers for Archiving are Changing

There have been many traditional drivers for retaining information for long periods in an archive, some of which are primarily strategic in nature and others that are more tactical or functional:

- Legal (eDiscovery, legal holds, early case assessment)
- Regulatory compliance
- Storage management and optimization
- End-user self-service to older content
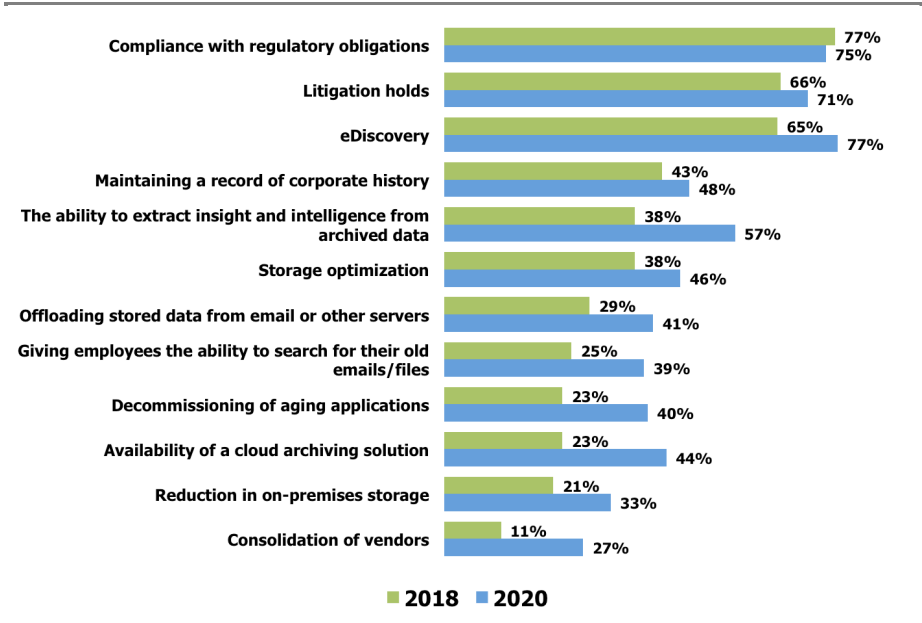- Retention of corporate "memory"
- Knowledge management

The survey confirmed that the most important drivers today are compliance with regulatory obligations, litigation holds and eDiscovery, as shown in Figure 3. Due to the frequency of lawsuits and compliance audits, enterprises need a way to preserve all types of relevant electronic content and they need to be able to search it quickly. Moreover, we found that these drivers are becoming more important over time.

....any electronic source of information... should be archived if it contains business information, or information that might be needed at a future date.

Compliance with regulatory obligations — 77% (2018), 75% (2020)
Litigation holds — 66% (2018), 71% (2020)
eDiscovery — 65% (2018), 77% (2020)
Maintaining a record of corporate history — 43% (2018), 48% (2020)
The ability to extract insight and intelligence from archived data — 38% (2018), 57% (2020)
Storage optimization — 38% (2018), 46% (2020)
Offloading stored data from email or other servers — 29% (2018), 41% (2020)
Giving employees the ability to search for their old emails/files — 25% (2018), 39% (2020)
Decommissioning of aging applications — 23% (2018), 40% (2020)
Availability of a cloud archiving solution — 23% (2018), 44% (2020)
Reduction in on-premises storage — 21% (2018), 33% (2020)
Consolidation of vendors — 11% (2018), 27% (2020)

■ 2018 ■ 2020

Source: Osterman Research, Inc.

The survey found that the ability to extract insight and intelligence from archived data is quickly growing in importance. The rapid rise in interest in business intelligence is a very important trend and is discussed in more detail later in this report.

The survey found that archiving remains an important content management tool in general. Storage optimization, offloading stored data from email servers, decommissioning aging applications, reduction in on-premises storage and consolidation of vendors are important content lifecycle management tasks served by archiving.

Storage optimization is an issue for email server performance and offloading inactive data, albeit primarily with on-premises systems. Old email and files that remain on production servers add unnecessary load to the application server and the backup server. It is best practice to regularly archive old and inactive files to a dedicated archive.

CLOUD ADOPTION
Availability of a cloud archiving solution importantly makes a jump in importance from 23 percent finding it an important or major driver to 43 percent in just two years' time. New cloud archiving solutions leverage low-cost cloud storage, and so as compared to traditional on-premises storage, cloud storage is available for a fraction of the total cost of on-premises solutions.

Cloud adoption in general is increasing rapidly. In the survey, respondents indicated that today 32 percent of their total enterprise content is stored in the cloud, but within two years 52 percent of total enterprise content will be stored in the cloud – a 62 percent increase in the space of just 24 months. This new trend is changing the notion of archiving in important ways.

Cloud storage provides built-in high availability so that the archived data is always protected. These reasons, plus the fact that employees can generally access the cloud using GUI-based applications to search old email and files, means that cloud storage is now the most popular target for long-term retention of electronic content.

Cloud adoption is causing organization to think about where content "lives" as opposed to the notion of transferring all content to a traditional, centralized archive. Examples include maintaining data in an SAP system or within Salesforce, as opposed to copying data in these systems to a separate, centralized archive. That's not to say that this content should not be archived using a traditional archiving approach on-premises or in the cloud, but decision makers are increasingly focused on managing the content as opposed only to storing it in an archive.

A MISCONCEPTION ABOUT THE CLOUD

There is a false impression that email data created and stored in cloud ecosystems like Office 365 doesn't need to be backed up beyond the native offerings, but that simply isn't true. For example, Office 365 does not offer traditional backup and recovery capabilities in the same way as organizations have deployed in on-premises environments. Office 365 is a live production system that offers recovery of messages and documents within a rolling time window. Instead, Microsoft uses alternative approaches for safeguarding current production data. For example:

- In Exchange Online, a user can recover a deleted item for up to 14 days by default (although an administrator can increase the recovery window to a maximum of 30 days).

- Data that is sent to the recycling bin from OneDrive will still be recoverable for 90 days, but only the most recent version of that data.

- Office 365 permanently deletes content from inactive or deprovisioned licenses.

In short, the use of a cloud-based email solution does not supersede an organization's responsibility to maintain a backup of its data. Backup is a responsibility of the customer, not the cloud provider.

Key Drivers for Content Services

For organizations that deploy content services, the survey asked respondents to rank the current and future drivers for content services. The survey results indicate that the leading drivers for deploying a content services solution are the ability to extract insight and intelligence, compliance with regulatory obligations and eDiscovery, as shown in Figure 4 on the next page.

Overall, each of the drivers shows a marked increase in importance as compared from 2018 to 2020. This result reflects the overall challenge facing all organizations as the number of content types increases, application data moves to the cloud and legal and regulatory pressures increase.
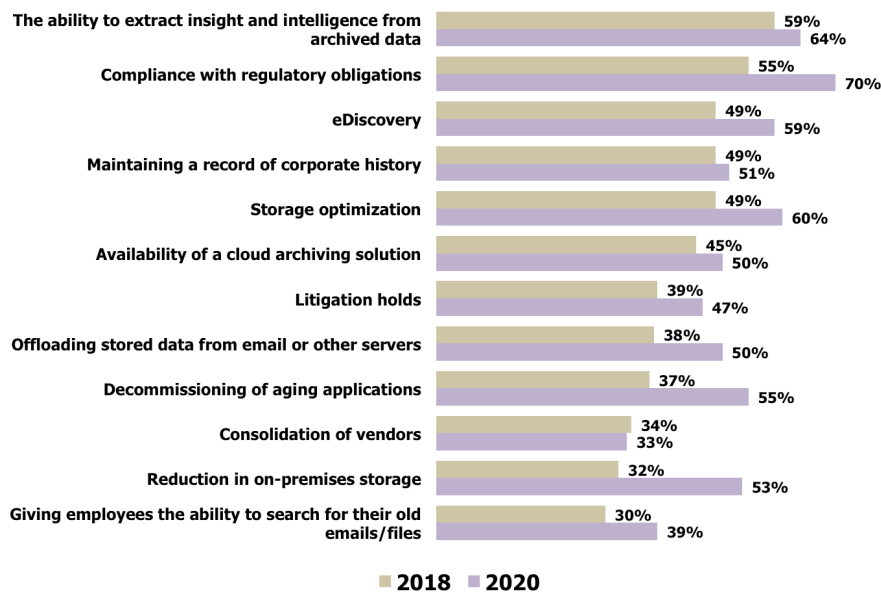
....the use of a cloud-based email solution does not supersede an organization's responsibility to maintain a backup of its data.

Figure 4
Drivers for Deploying a Content Services Solution, 2018-2020
Percentage Responding an Important or Major Driver

Driver	2018	2020
The ability to extract insight and intelligence from archived data	59%	64%
Compliance with regulatory obligations	55%	70%
eDiscovery	49%	59%
Maintaining a record of corporate history	49%	51%
Storage optimization	49%	60%
Availability of a cloud archiving solution	45%	50%
Litigation holds	39%	47%
Offloading stored data from email or other servers	38%	50%
Decommissioning of aging applications	37%	55%
Consolidation of vendors	34%	33%
Reduction in on-premises storage	32%	53%
Giving employees the ability to search for their old emails/files	30%	39%

■ **2018** ■ **2020**

Source: Osterman Research, Inc.

Data Types for Content Services

Content services manage a wide range of content data types, both on premises and in the cloud. In the survey, we asked organizations that deploy content services to rank the importance of a wide range of content types. The results shown in Figure 5 reveal that email ranks number one, followed by user files, content from SharePoint or similar collaboration tools, and CRM data.

Figure 5
Data Types That are Important to Include as Part of a Content Services Solution

Content Type	%
Corporate email	82%
Users' files	77%
Content from SharePoint or similar collaboration tools	68%
CRM data	66%
Content from company-managed file sync and share tools, e.g., Dropbox or OneDrive	59%
Content from departmental file shares, e.g., Legal, Finance	52%
Accounts payable content	45%
Content from image capture systems	41%
SAP data	41%
Content from company surveillance video	39%
Content from company-owned mobile devices	36%
Voicemails from the company phone system	36%
Corporate instant messaging/mobile message app content	34%
Content from ex-employees	30%
Corporate Facebook posts	27%
Corporate LinkedIn posts	27%

Content Type	%
Content from users' personally managed file sync and share tools	25%
Text messages	25%
Work-related content from employees' personal mobile devices	25%
Corporate Twitter content	23%
Work-related content from employees' instant messaging accounts	23%
Work-related posts from employees' personal social media accounts	20%
Other	16%

Source: Osterman Research, Inc.

The content that organizations consider worth retention for long periods of time (or that they should retain for long periods) contains a wealth of valuable information. For example, consider what is contained in email as just a single source of many different content types that should be retained:

- A record of every message sent between employees, business partners, customers, prospects, contractors, suppliers and others.

- The way employees deal with confidential information. Even if data loss prevention or encryption technologies are employed and will block the inadvertent transfer of this information in violation of corporate policies, it is useful to know which employees are ignoring corporate policies or need more training on them.

- Any messages and files that employees send to competitors.

- A second-by-second record of employee activities during working hours and when they are away from the office working on organizational business.

- A detailed timeline of each employee's responses to time-sensitive messages.

- A "sort-of" transcript of how employees respond to the organization's customers and prospects, as well as the messages that employees ignored or never saw.

- A record of all of the business relationships maintained by employees.

- A detailed record of every calendar appointment set by employees, their location, and other relevant data about how they spend their day.

A thorough analysis of this content can reveal a great deal of information that decision makers can use to analyze their operations more completely. For example, an in-depth analysis of email content can reveal:

- How employees comply (or don't comply) with corporate policies about protecting confidential information or the organization's email acceptable use policies.

- How sales staff interact with the organization's customers and prospects, such as how quickly they respond to customer inquiries.

- Each of the social relationships that exist within a company, which might help fellow employees find the expertise and influencers they need to work more effectively.

- Corporate sentiment about employees, managers and the organization overall.

A thorough analysis of…content can reveal a great deal of information that decision makers can use to analyze their operations more completely.

Using this information, decision makers can more thoroughly understand things like:

- Who the key players in an organization might be and how that differs from what the organization chart might reveal.

- How and to whom information flows inside and outside the organization.

For example, an email archive or email system contains the actual emails between customer service representatives and customers. However, it also reveals how quickly these representatives responded to customer inquiries, the tone of their responses, any responses from customers, and a wealth of other information. This can be valuable in more thoroughly understanding how customers are treated and how likely they might be to renew or expand their relationship with a vendor. This data can be correlated with data from CRM systems, social media platforms, and even text messaging systems to provide a more thorough understanding of the sales and support process.

In short, today's defensive approach to archiving and content management stores information – the next generation of archiving and content management focuses on understanding what is contained in that information.

Use Cases

To understand use cases for content management and archiving, we did not want to restrict respondents to a set of answers, but instead allowed them to answer freely. The responses were varied, but there emerged a set of responses around a need for gathering business intelligence from electronic content to help business management. Decision making, meeting compliance requirements, user behavior and business process performance were commonly identified as important drivers for content management:

- Meet all compliance requirements

- Gather actionable, accurate, and timely data

- Provide insight to management to enable more informed decisions

- Identify risky behaviors, hold data-led conversations on areas of concern, and put together necessary remedial plans

- Faster and better decision making for enterprise and regulatory compliance

- Analyze usage patterns, user behavior, and corporate policy violations

- Analyze trends that are longer-term in nature, such as possible missed opportunities

- Develop retention strategies

- Analyze information and how it is used in day-to-day operations

- Glean specific intelligence accurately

- Expertise location, coordination of efforts, cross-sell opportunities

- Patterns in outages, process delays, implementation schedules and velocity

IMPROVING THE SALES AND SUPPORT PROCESS

A properly implemented and managed CRM solution can contribute significantly to the sales process in an organization. However, not all relevant sales information is

captured within CRM systems. There is a great deal of data that cannot practically be entered by salespeople about customer and prospect interactions, such as the tone of a customer's inquiry, the number of minutes between receipt of an inquiry and responding to it, and the various resources that were used to make a sale. However, if information from other content sources can be run through an analytics solution, it can answer a variety of important questions about an organization's sales process. For example:

- How long does it take for the customer or prospect to respond to the salesperson's response and does that correlate in any way with sales success or customer retention rates?

- Are sales teams focused on the right opportunities?

- What is the length of time required for a member of the sales staff to respond to a customer's inquiry or complaint?

- Does it take longer to respond to a complaint compared to an inquiry?

- Do salespeople loop in others to help answer inquiries and does this help in the sales process?

- What is the correlation between the sentiment that salespeople use in email or other communication modes and the likelihood that a customer or prospect will buy?

- Is the mode of communication that a sales staff member uses correlated with sales success?

BRINGING EMPLOYEES UP TO SPEED MORE QUICKLY

New employees rely on information to come up to speed quickly in their position. In a typical hiring scenario, a new employee is provided with a fresh email account and a network file share. The challenge for the new hire is accessing all of the electronic information they need to learn the new role. Existing files, email and other electronic content can be spread across file shares, content management systems and inactive email accounts. Some of this data may or may not be accessible given the highly siloed nature of most content today. The right content management and archiving solution provides a single repository for long-term retention of electronic content, or the tools with which a new hire can easily access information from multiple siloes.

With the right tools, new hires are provided with archive/content access and the tools to search across thousands of documents to find the information they need. New employees accelerate the learning process by accessing email communications with customers and suppliers, reports, presentations, spreadsheets and more. Additionally, access to social media content provides insight into market conditions and customer sentiment.

Content analytics gives new employees a tool to search across all forms of content for specific topics. The benefit is that the new employee can target information he or she needs and not have to wade through thousands of pages of unhelpful information.

PRESERVING CONTENT WHEN EMPLOYEES LEAVE

When employees leave, or when companies merge or are acquired, a major issue is how to manage all of the electronic content for these employees. Examples of relevant electronic content that should be retained and accessible are email, instant messaging data, social media data and files. One approach, commonly employed by many organizations, is to do nothing: simply disable access for the departed employee and leave all of his or her data in place. However, there are two problems with this approach:

Content analytics gives new employees a tool to search across all forms of content for specific topics.

- First, the data remains in isolated application silos, making it cumbersome to search, if it is searchable at all.

- Second, the data continues to take up valuable primary storage space while yielding little to no benefit to the organization because it cannot be accessed easily.

A better solution is to consolidate personal data from departed employees into a dedicated archive or content management solution. Special archive connectors are used to connect with the data siloes (e.g., email, instant messaging, files and social media). Once data is moved into the archive, the data is secured for regulatory and legal purposes, and it can be searched quickly for purposes of eDiscovery and audits. After the inactive data is moved, space is freed up on primary storage, creating room for new information.

THE ABILITY TO FIND EMPLOYEES WHO WILL CAUSE A DATA BREACH

Research has discovered that poorly treated employees are more likely to commit fraud against their employers by sending sensitive or confidential information to competitors, stealing data or committing some type of financial fraud. An important way to discover how managers are treating employees is to analyze communications between managers and employees, as well as inter-employee communications. This might reveal that employees are treated poorly, and it could help senior managers to take the proactive steps necessary to prevent managers from acting inappropriately. This can significantly reduce corporate risk.

CONDUCTING INVESTIGATIONS

Resolving employee and customer disputes is a common challenge for all organizations. The most difficult challenge is getting a clear understanding of what took place and when it occurred. Too often, people do not have a clear recollection of what took place or what was said. Luckily, email and instant messages provide indisputable evidence of what was said (assuming that chain-of-custody was preserved). The result is that these sources are the most popular type of electronic content used in eDiscovery.

Today, electronic discovery of email is the number one driver for content archiving and the technology to index email and attachments is quite mature. Properly designed archiving and content management systems can search millions of emails in a short period of time. Besides indexing, other advanced eDiscovery features support custom tagging and categorization. For example, when email is tagged as "non-responsive", the need to search the same email again in a future legal matter is eliminated.

eDiscovery of email continues to be a key driver for email archiving – next generation archiving expands this capability by including social media, hosted applications and file-sharing applications as additional sources of information for legal, regulatory and analytics purposes.

ENABLING EMPLOYEES TO FIND OTHERS MORE EASILY

This use case is related to onboarding new employees, but it also pertains to all employees, so it will be discussed separately. One challenge all employees face is accessing the information they need as quickly as possible, and many times the source of key information is another employee. Good electronic content analysis systems can enable employees to find subject-matter experts easily and quickly.

Using content analytics, employees can search by keyword to identify matching content contained in email, blogs, reports, CRM and more. For email and other forms of content, the content owner is contained in the content metadata. A simple content analysis could rank search results by content author. Authors who publish most frequently can quickly be identified.

The same use case works for customer support. A simple content search analysis can identify all of the employees (e.g., those in sales or support roles) who have had contact with a customer. This offers a significant benefit when an employee needs to understand a customer's history quickly, perhaps in real time.

POLICY AND LEGAL COMPLIANCE

Good analytics can be used to identify serious problems so that violations of corporate policy, the law, or best practice can be addressed before they result in a serious problem. For example, a company's compliance staff could search for evidence of sexual harassment, lying to customers, illegal downloads, distribution of offensive content, or any of several other activities that could result in a regulatory action, lawsuit, scandal or some other problem.

UNDERSTANDING WHAT EMPLOYEES ARE TALKING ABOUT INFORMALLY

Good analytics can help decision makers understand what employees are talking about on the job. Some employees talk at length about corporate rumors like layoffs in email or text messaging, which can lead to poor morale and increased employee turnover. If managers can identify these rumors and their source, and understand why employees are concerned about what is happening in the organization, this might help senior managers to communicate more effectively and address problems as early as possible.

THE EUROPEAN GENERAL DATA PROTECTION REGULATION

The European General Data Protection Regulation (GDPR) imposes a major burden on any organization that has data on residents of the European Union (EU). The GDPR will demand a level of data retention and management that is at least equivalent with the level of effort required to properly manage eDiscovery activities. Moreover, these activities must normally be performed at no charge, and so content management must be efficient and easy to use. There are some important implications to consider for organizations that possess data on residents of the EU, even for organizations that do not have any operations in the EU. For example:

- Article 15 of the GDPR gives data subjects the right to ask any entity that controls or processes his or her personal data to provide that data on demand. EU data subjects also have the right to know if and when their data is transferred to another country or organization, along with whatever safeguards that are in place to ensure the protection of their data after it has been transferred. A data controller must provide a copy of any personal data that is being processed at no charge the first time it is requested and they have only one month to do so (three months in exceptional cases).

- Article 17 provides that, subject to certain conditions, a data subject has the "right to be forgotten" by any data controller that possesses or controls his or her information.

- Article 30 provides that data controllers must maintain records of their data processing activities, with a list of specific information to be retained for each record.

As a result, organizations must retain and properly manage data like spreadsheets and other data sources populated by exporting customer contact and profiling details for a mail merge; copies of production databases that contain personal data taken for testing, development, or analytics purposes; and a wide variety of other data types. In short, GDPR will require the deployment of very robust archiving and content services solutions.

The GDPR will demand a level of data retention and management that is at least equivalent with the level of effort required to properly manage eDiscovery activities.

Archiving and Content Services Challenges

Many current archiving and content services solutions were not designed to be proactive. Electronic content, in all its forms, contains a record of business. Decision makers need to be able to use this information proactively to extract insight and intelligence about the organizations and its external relationships, to gain additional understanding about information flows that are not captured in other corporate systems, and to manage risk more effectively. Traditional, legacy archiving solutions present problems when organizations are trying to implement next-generation archiving and content services capabilities:

- **Data is siloed**
  Content is normally stored in many independent silos, such as email, CRM or SAP. Because of the growth of corporate applications, particularly those in the cloud, and IT's growing acceptance of the "Bring Your Own" trend for applications and devices, information management is becoming highly fragmented and distributed. This makes data increasingly difficult to access by senior managers, legal teams, compliance and others who must access it. Ad hoc data inquiries, such as those that individual employees might initiate, are more difficult if an organization is storing a variety of data types. The result for end users is that they often do not look for older data because of the time and difficulty involved in doing so.

- **Data in different siloes is not connected**
  Corporate information is increasingly distributed and not connected, which leads to an inability to search and synthesize easily across the various siloes in which it is stored. This means that information access requires visiting a number of data siloes one at a time, such as CRM, email, social media, etc., each with its own interface. Plus, data cannot be connected between siloes in most cases.

- **Siloes lead to increased risk**
  When organizations are faced with an eDiscovery order, a regulatory audit, an early case assessment or even just an informal search of information, they must search across multiple siloes, as noted above. The fact that all relevant data is not accessible in a single interface increases the likelihood that important information will be missed – and the individual or group searching for the information may never know it. That can significantly increase corporate risk from a legal, regulatory or productivity standpoint.

Next generation archiving and content services solutions, on the other hand, are designed with analytics in mind and can traverse across all forms of electronic content, regardless of location. In the previous section, we discussed several ways in which electronic content analysis can benefit the overall organization. Using analytics for business intelligence, new content services can analyze content in on-premises file shares, hosted email platforms and hosted file-sharing systems, all in a single search. Improved access to content and improved analytics are two key enablers for next generation archiving.

Next Steps

Osterman Research recommends a four-step approach to developing a next-generation approach to archiving and content services:

1. **Focus on archiving as a proactive tool to safeguard business data, empower data extractors, and glean actionable business intelligence**
   First, understand that the traditional role of archiving is primarily a defensive one. Any organization that has faced a major eDiscovery or a regulatory audit knows very well how costly it can be to search and find electronic content (e.g.,

email, files, social media) when it is not already preserved in a dedicated archive. The cost of a single legal eDiscovery case can more than pay for the cost of an email archiving solution.

Yet, today many organizations still do not archive their content. It is recommended that these organizations look at email archiving as a first step. The good news is that new cloud-based enterprise email solutions (e.g., Microsoft Office 365 or Google Gmail) support built-in archiving, and so a separate, third party email solution is not an absolute requirement in all situations. While Osterman Research highly recommends the use of third-party archiving solutions, they are not an absolute requirement to get started on the road to a basic archiving capability.

2. **Decide how data should be retained**
Second, decision makers should decide how all their data should be retained and managed. One option is to collect copies of archived data into a central repository and maintain it under IT control in a single archive. While this is feasible and can be a good solution in many cases, it can cause problems. For example, there could be a major increase in storage requirements because the central archive will now contain a duplicate copy of archived content that is also stored in other locations. At least some of the data in the central archive will be out of sync with data in the original archives as changes are made to the latter between replication/syncing cycles. Plus, not every piece of content needs to be archived, and so copying all data to centralized archives will store large quantities of unnecessary data. This adds to storage and storage management costs.

A better approach might be to implement a solution that will enable retention and analysis of content "in-place" rather than moving it into a centralized archive. This can reduce storage requirements, can make data management much easier, and allow easier integration of legacy data stores into the overall content repository.

3. **Seek advantage in extracting insight and intelligence from data**
The ability to extract information from retained content, whatever its source, is incredibly valuable and can provide insight available nowhere else. Performing analytics on this data can offer significant competitive advantage for any organization and may allow them to gain new insights from their data stores. However, decision makers must understand these benefits and authorize the effort necessary to realize them. Getting buy-in for these kinds of solutions is not an easy thing to do in some organizations, especially in those that are not yet sold on the concept of traditional archiving.

4. **Sell it on the use cases**
The use cases presented earlier provide examples of how next-generation archiving can help drive better decision-making and enable greater employee productivity. One key is to use analytics for business intelligence to gain a competitive advantage over organizations that are not mining their content for business insight. Content services solutions enable secure access to relevant information across content sources that would otherwise go unleveraged.

Summary and Conclusions

Traditional archiving and content management is an essential best practice for any organization, since it permits organizations to mitigate their legal, regulatory and other risks. However, content stores contain a wealth of information about how the organization really operates, how information flows within an organization, employee social relationships, and other valuable information. The application of content analytics to content stores can deliver actionable results and competitive advantage in a way that other solutions cannot. It can enable organizations to use their rich

The ability to extract information from retained content, whatever its source, is incredibly valuable and can provide insight available nowhere else.

stores of information to proactively identify and address problems before they get out of hand, rather than purely for discovery after the fact. Every organization should seriously consider deploying a next-generation archiving or content services solution that will enable it to extract insight and intelligence from its information stores.

Sponsor of This White Paper

Viewpointe helps to address information challenges and automate content processes with OnPointe, a suite of content service solutions that help improve information management and document processes across the enterprise with the flexibility to focus on current content challenges while also providing a foundation to address future business needs. Viewpointe has the expertise to assist with professional services to guide implementation and project success, content services to meet functional requirements with system connectivity and embedded end-user interfaces and proven managed services in a private cloud environment to keep content secure. In addition, Viewpointe delivers the promise of flexible, future-proofed cloud services without sacrificing the security and compliance demands of regulated businesses.

Viewpointe allows you to take control of business content across a broad range of enterprise applications, messaging systems, content repositories and even file shares that often go unmanaged. Designed to handle petabytes of information, OnPointe provides specific services and integrations to ingest and classify many types of structured and unstructured information. With our extensive APIs, even custom or in-house developed applications can be integrated, helping to insure consistent governance of data.

Five key capabilities comprise the OnPointe suite of private cloud-based content services include:

- Content capture and management
- Search, retrieval and version tracking
- Security and privacy controls
- Retention management and disposition
- Workflow automation of content processes

Viewpointe's cloud-based managed services help streamline the information lifecycle and automate business processes with a foundation of compliance and information governance. Known for its time-proven expertise in information strategies and managed services, Viewpointe makes it possible to address today's business opportunities and risks while planning for tomorrow's challenges. Founded in 2000, Viewpointe has been named to the prestigious FinTech 100, representing the best in financial technology and service providers every year since 2006 and is the chosen partner at many of the world's most esteemed companies.

Viewpoint*e*

www.viewpointe.com

@Viewpointellc

+1 800 956 3807