Archiving as a Key Element of Good Information Governance

White Paper by Osterman Research
Published June 2020
Sponsored by OpenText

Contents

Archiving business records is the first step in enabling a proper defense during regulatory audits, legal actions, and in other situations in which older content must be retained.

Executive summary

Information governance can be defined as "policy-based control of information to meet all legal, regulatory, risk, and business demands"[1]. It is an essential best practice for any organization, but particularly larger ones that must satisfy a complex set of regulatory, legal and other compliance demands. Archiving technologies – solutions designed to capture, index and store content – are an integral component of any information governance capability and must be considered as part of an overall information governance initiative.

Key takeaways

- Proper governance of information is becoming more difficult to address over time because of increasing volumes of information, a growing number of regulatory requirements, more stringent legal requirements to retain and produce data, and the need to defend decisions to delete information that is no longer necessary to retain.

- However, many organizations don't have an information governance program. As shown in Figure 1, about two-thirds of mid-sized and large organizations have an information governance program, but one-third do not. Most of the remaining organizations plan to implement such a program by mid-2021.
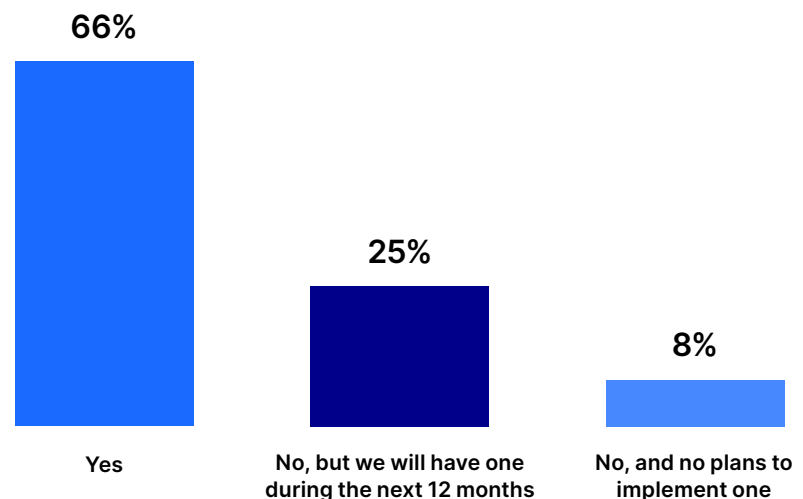

Figure 1 "Does your organization have an information governance program?"
Source: Osterman Research, Inc.

- Archiving business records is the first step in enabling a proper defense during regulatory audits, legal actions, and in other situations in which older content must be retained. However, the same set of archived data can be used more proactively to glean insights and intelligence about how a company operates, what its employees are doing on a daily basis, and whether or not policies are being followed.

- Good archiving technology is an essential element of any information governance strategy and for extracting insights and intelligence about a business, and it's key to managing information properly. However, many organizations are not pleased with their current archiving solutions.

1 https://community.aiim.org/blogs/robert-smallwood/2014/08/18/defining-the-differences-between-information-governance-it-governance--data-governance

To ensure that information governance is effective, a robust good archiving capability should be implemented.

About this white paper

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

Information Governance is a Challenge

Information is vital for almost any organization, and so proper governance of that information is essential. Proper information governance can offer significant benefits in a number of areas, such as reducing the costs of regulatory compliance, eDiscovery and litigation support; it can reduce corporate risk; and it can improve employee productivity by making it easier and faster for employees to search for and find necessary information. The return-on-investment associated with good information governance can be enormous for most organizations and it should be considered a high priority.

To ensure that information governance is effective, a robust good archiving capability should be implemented to ensure that the appropriate information is being captured, indexed, stored and made available as efficiently and quickly as possible. In short, good information governance that has at its core robust archiving functionality can address a number of critical issues that face every organization:

- **Managing the growing flood of information**
Information grows exponentially and can overwhelm those attempting to manage it. Even if an organization retains only its email, over several years an organization can generate tens of millions of emails that must be archived and properly managed.

- **eDiscovery and litigation support**
Eventually, almost every organization will be involved in litigation in some capacity, and they will need a reliable and efficient way to search for all of the relevant information that might be needed. In the absence of a robust archiving capability, eDiscovery, early case assessments and other types of litigation support will be extremely difficult and expensive, if not impossible.

- **Managing regulatory compliance**
A large proportion of organizations must satisfy a rigorous set of compliance obligations, and so must govern their information properly. However, in an era of privacy regulations like the General Data Protection Regulation (GDPR) or the California Consumer Privacy Act, there are no "unregulated" organizations – most every organization is subject to some form of requirement to retain and produce information. Good information governance in conjunction with robust archiving capabilities are essential to satisfy these regulatory requirements.

- **Maintaining user productivity**
Users search for information on a regular basis to do their work. The survey conducted for this white paper found that during a typical week, users spend an average of 71 minutes searching for information that already exists. Over the course of a 50-week workyear, that amounts to just over 59 hours each year that users are spending searching for information. If we conservatively estimate that a good archiving solution as part of a good information governance program could halve that figure, that would return 3.7 days of productive work each year to each information worker.

An essential element of good information governance is being able to dispose of information that is no longer needed, and doing so in a way that can be defended to regulators and the courts.

- **Managing storage bloat**
  Storage bloat is one of the primary problems that organizations face in the context of growing information stores. Even in an era of "unlimited" cloud storage, bloat still manifests itself in a variety of ways, such as higher storage costs, greater difficulty in finding needed information for eDiscovery, producing too much information when only a fraction of it is needed, and so forth. Good information governance and archiving are useful capabilities in eliminating unnecessary information and minimizing storage bloat.

- **Defensibly disposing of information**
  As a corollary to the point above, an essential element of good information governance is being able to dispose of information that is no longer needed, and doing so in a way that can be defended to regulators and the courts. Defensibly deleting information reduces corporate risk and lowers storage management costs.

- **Records management and retention practices**
  At the core of good information governance is being able to manage business records properly and retaining information for the appropriate lengths of time. However, many organizations don't really know for how long they are to retain their information: the survey conducted for this white paper revealed that only 25 percent of those involved in information governance for their organizations know exactly what to retain and for how long they should retain it; another 14 percent are not sure what to retain or for how long it should be kept.

In short, information governance is fraught with challenges that must be addressed if information is to be managed properly.

Why Organizations Must Archive Electronic Content

There are a number of reasons that organizations must retain their electronic content in an archiving platform. However, our research finds that those involved with information governance issues are not all that pleased with their current archiving solutions. For example, as shown in Figure 2, only 62 percent agree that their current archiving solution meets their regulatory needs, and only slightly more than one-half believe that their current solution is cost effective.

Meets our need to retain data for regulatory purposes	62%
Is cost-effective	53%
Meets our needs overall	51%
Retains all of the content we need to retain	50%
Retains all of the content we would like to retain	50%
Meets our eDiscovery needs	46%
Meets our need to do early case assessments	42%

Figure 2 Agreement With Various Statements About the Current Archiving Solution
Percentage indicating they "agree" or "strongly"
Source: Osterman Research, Inc.

The proper management of eDiscovery capabilities in conjunction with a good archiving solution is a best practice – relying on backups to do so is not.

This section discusses the key reasons that archiving is an essential element of any information governance program.

Archiving is a good defense

The traditional role of archiving is defensive: retain, and be able to access, business records in the event they are needed to satisfy a regulatory audit, an eDiscovery order, inform senior managers' and legal counsel's early case assessments, and the like. The goal of archiving as a defensive tool is to ensure that information is retained in case someone comes looking for it. While the defensive aspect of archiving will never go away, archiving is increasingly being used for more proactive purposes, as will be discussed in the next section, Extracting Insight and Intelligence From Your Archives.

Legal considerations

All organizations must retain various types of electronic content in the event this content is needed to support their role as a defendant, plaintiff or third-party participant in a legal proceeding.

The need to retain and manage information is imposed by statutes, legal precedent, and internal best practices. Further, organizations that reasonably expect upcoming litigation are obligated to place relevant electronic content on legal hold that may be different from their standard policies, even if they have not been informed of pending litigation. An archiving solution can facilitate that process in a way that traditional backup processes cannot.

The proper management of eDiscovery capabilities in conjunction with a good archiving solution is a best practice – relying on backups to do so is not. Using an archiving solution, information is easier and faster to find than it is when using just a backup, the data set that is captured is more complete (which reduces the risk of problems like spoliation), and it's normally less expensive because less work is required to gather and produce the data. Moreover, having an archive with proper eDiscovery functionality also helps organizations to bring much of their eDiscovery work back in-house, and it allows them much faster early case assessment, particularly for smaller cases.

Regulatory compliance

A large volume of the electronic records related to an organization's activities are subject to a variety of regulatory obligations, particularly in heavily regulated organizations like those in the financial services, healthcare, energy and life sciences industries. However, it's important for decision makers to realize that almost every organization and industry faces some level of regulatory compliance obligation to retain records, and that these obligations are not limited merely to a few industries. As just a few examples:

- **Healthcare**
  The Health Insurance Portability and Accountability Act of 1996 (HIPAA) and various other obligations mandate that healthcare organizations must protect patients' electronic health records from unauthorized users and to retain this information for at least several years, sometimes indefinitely. Non-compliance with these requirements can result in significant fines and, in some cases, prison sentences for the most egregious offenders.

Good archiving enables greater user productivity by allowing employees and others to spend less time looking for content.

- **Financial services**
  In the United States, a variety of regulations from the Securities and Exchange Commission and the Financial Industry Regulatory Authority require brokers and dealers, registered investment advisors, financial advisors and others to retain various transaction records for approximately six years. In the United Kingdom, investment service and transaction records need to be retained for at least five years as required by the Financial Conduct Authority. In Canada, records of securities purchase and sell orders must be retained for seven years as required by the Investment Industry Regulatory Organization of Canada. The European Union has similar requirements under the Markets in Financial Instruments Directive. A failure to comply with these obligations can result in sanctions, fines and various other penalties.

- **Nuclear materials**
  Entities that house and transport various types of nuclear materials must retain records for three years about authorized individuals, visitors, vendors and others; routine security tours and inspections; alarm events; shipments; procedures and other information.

- **Air transportation**
  Every airline and other air transportation companies offering passenger services operating to, from or within the United States with at least one plane that has 30 or more seats is obligated to keep records for two years about tarmac delays that last more than three hours, including the reason for the delay, the length of the delay, the actions taken to minimize hardships for passengers, the disposition of the delayed flight, and an explanation for any delay in excess of three hours.

- **Chemicals**
  Companies that are involved in the chemicals industry must retain investigation reports for five years about incidents in which a serious release of a regulated substance occurred or might have occurred.

These are just a few examples – there are literally thousands of data retention requirements across virtually every industry.

End-user self-service

Information workers must periodically find old documents, presentations, spreadsheets and other information in the normal course of doing their work. As noted previously, the average information worker spends 71 minutes per week doing so. Plus, new employees will often need to refer to communications and other content generated by people who held their jobs previously. An archive is an essential tool in helping employees to search and find old information instead of spending inordinate amounts of time searching for it, or worse, recreating it. Good archiving enables greater user productivity by allowing employees and others to spend less time looking for content.

Not archiving all necessary business records during the period of the COVID-19 crisis is going to have longterm ramifications for businesses that fail to retain and protect their business records properly.

Storage management

While storage management is not quite as important today as it was when storage was more costly, one of the benefits of a good archiving solution is its ability to improve system performance by reducing the amount of "live" data that must be stored. While older data must still be retained, sometimes indefinitely, it is used less frequently as it gets older and does not need to be retained in active storage. Migrating this data to an archive will improve system performance, and will also speed the process of backing up and restoring servers and workstations after system crashes or other problems. Another key benefit of archiving in the context of storage management is its ability to deduplicate and provide single-instance storage, which offers the potential to significantly reduce storage requirements.

Knowledge management

Finally, a good archiving solution as part of a robust information governance program will enable good knowledge management – the retention and management of information that may be useful for activities outside of supporting regulatory compliance, litigation support and other external drivers.

The impact of the pandemic

The COVID-19 pandemic has forced tens of millions of information workers into a work-from-home environment, a scenario that has caught many IT, security and information governance teams off-guard. Reflecting the new reality is that many organizations are no longer archiving their data properly. As shown in Figure 3, an April 2020 Osterman Research survey found that only 59 percent of organizations are archiving data as they were before the work-from-home phenomenon began, and 10 percent are no longer archiving their data at all.
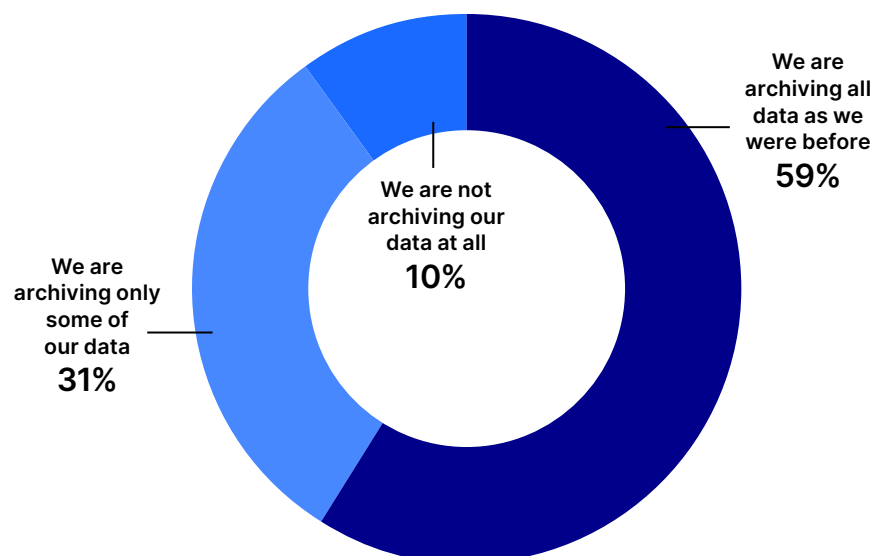

We are archiving all data as we were before
59%

We are not archiving our data at all
10%

We are archiving only some of our data
31%

Figure 3 Archiving Practices During the COVID-19 Crisis
Source: Osterman Research, Inc.

Both compliance and security capabilities are now reduced during the work-from-home phenomenon.

This underscores a serious compliance issue for many organizations. Even though employees are working from home, that in no way does away with corporate requirements to satisfy compliance obligations. Not archiving all necessary business records during the period of the COVID-19 crisis is going to have long-term ramifications for businesses that fail to retain and protect their business records properly.

Inadequate archiving leads to inadequate compliance

All organizations should archive their business records for a variety of reasons, including regulatory compliance, litigation support, eDiscovery, early case assessments, enabling users to find older information, and so forth. Even organizations that are outside of heavily regulated industries, such as financial services or healthcare, must archive content to comply with a growing body of privacy regulations like the European Union's General Data Protection Regulation and the California Consumer Privacy Act.

However, Osterman Research data from numerous surveys, including the one conducted for this white paper, found that not all organizations are archiving their business records. What makes the problem worse, as implied in the figure above, is that the COVID-19 pandemic is prompting even fewer organizations to archive their business records, potentially (and very likely) creating more compliance and legal problems in the figure.

Figure 4 shows that both compliance and security capabilities are now reduced during the work-from-home phenomenon. For example, while 64 percent of organizations report they agree with the idea that they were doing an excellent at job maintaining compliance with their various obligations before the COVID-19 crisis, that level of agreement has fallen to 56 percent now. Similarly, while 56 percent were in agreement that they were doing an excellent job at blocking security threats before, that figure has dropped to 49 percent today. Even under normal circumstances figures for doing an excellent job at compliance and security should not be this low, but the current crisis has had a significant impact on organizations' ability to deal effectively with both. The reduced ability to archive business records during the pandemic is playing a major role in the reduced ability to meeting compliance obligations.

Issue	%
Before this crisis, we were doing an excellent job at maintaining compliance with our various compliance obligations	64%
During this crisis, we are doing an excellent job at maintaining compliance with our various compliance obligations	56%
Before this crisis, we were doing an excellent job at blocking security threats from impacting our organization	56%
During this crisis, we are doing an excellent job at blocking security threats from impacting our organization	49%

Figure 4 Changes in Compliance and Security Posture From the COVID-19 Pandemic
Percentage responding "agree" or "strongly agree"
Source: Osterman Research, Inc.

Extracting insight and intelligence from your archives

As discussed in the previous section, the traditional and primary role of an archiving solution has been mainly defensive in nature: keep records in the event they are needed by an external party like a court or regulator. However, could organizations archive data not only for defensive purposes, but also to take a more proactive approach and glean insights and intelligence from their archived data? The answer is definitely yes. Consider what just an email archive contains:

- An indexed, immutable and time-stamped copy of every communication and attachment sent by everyone with access to the corporate email system, including the senders and recipients of these emails.

- Full audit reports showing who has accessed what information, when they accessed it, etc.

- The wording and tone of every email that has been sent and received.

- A record of how quickly recipients respond to emails that call for some kind of response, such as a customer request or a prospect inquiry.

- A record of when during the day (or night) employees are conducting company business.

- Whether or not sensitive or confidential information was sent encrypted in compliance with corporate policies.

- Employees' views on clients, prospects, business partners, fellow employees, their managers, and others. It also includes employees' use of profanity, sending of racially or sexually offensive content, and potential instances of fraud.

- A record of employees or contractors using corporate email resources for personal uses.

What if decision makers analyzed their archived data to understand the timeliness with which salespeople respond to customer inquiries and how this translates to sales? What if they analyzed email data to determine if certain types of employee behavior correlate with a greater likelihood of data theft? What if they analyzed employee sentiment so they could better understand problems among employees before they could result in serious issues? What if they could monitor compliance with ethics standards and internal compliance rules to catch violators early on and thereby avoiding bigger problems later?

The key drivers for archiving

What drives the need for archiving solutions? As shown in Figure 5, the drivers are still focused primarily on defensive ones – regulatory compliance, litigation support and the like. Moreover, the ability to extract insight and intelligence from archived data is one of the least important drivers, cited by only about one-third of information governance professionals as important or more significant. However, while all of the drivers will become more important over the next two years, the importance of the proactive use of archiving will grow at a faster rate than others.

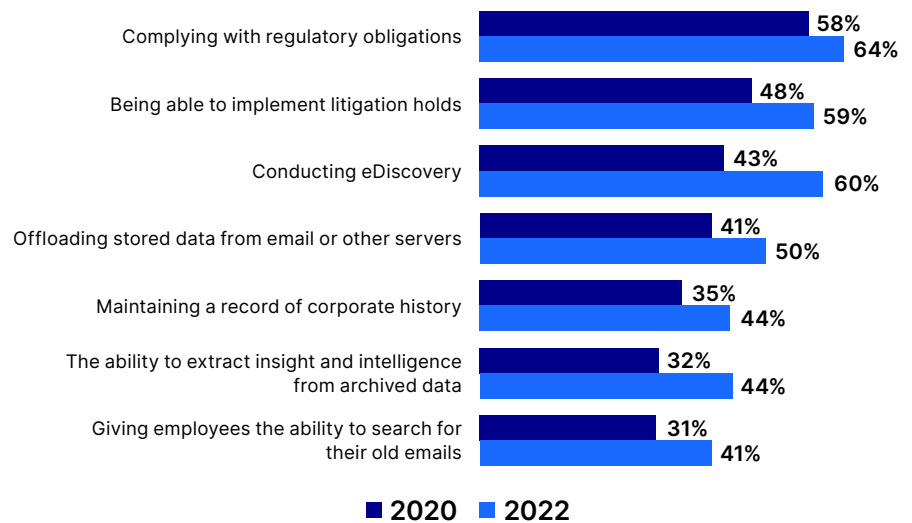**The key to determining what needs to be archived is the set of applications that generate business records.**

Driver	2020	2022
Complying with regulatory obligations	58%	64%
Being able to implement litigation holds	48%	59%
Conducting eDiscovery	43%	60%
Offloading stored data from email or other servers	41%	50%
Maintaining a record of corporate history	35%	44%
The ability to extract insight and intelligence from archived data	32%	44%
Giving employees the ability to search for their old emails	31%	41%

■ **2020** ■ **2022**

Figure 5 Drivers for Deploying an Archiving Solution, 2020 and 2022 Percentage responding an "important" or "major" drive
Source: Osterman Research, Inc.

Information archiving moving forward

Many organizations are well-served by an archiving solution that archives only email. However, additional information types will be archived increasingly in the future, such as text messages, social media posts, information from collaboration tools, video meetings, and the like.

The key to determining what needs to be archived is the set of applications that generate business records. Today, the vast majority of business communications, and even collaboration, takes place in corporate email, and so it makes sense for organizations' archiving implementations to begin with email. As organizations generate more business records and more archivable content in other platforms and applications, it makes sense to consider adding that content to the archive. An archiving solution that, in addition to email, retains all relevant data types, such as text messages, tweets, social media posts, instant messages, content from collaboration systems, and other data can be used to gain even more insight about how the organization runs.

Extracting insight and intelligence can yield enormous benefits

It's clear that a properly configured and managed archiving solution can yield significant benefits when used solely for defensive purposes. It can allow timely response to regulators' requests for information, to legal inquiries, and the like. However, the ability to conduct proactive, in-depth investigations is also enabled by use of a robust archiving capability. Thorough investigations require gaining understanding of what took place and when. Because individuals often do not have a complete recollection of what took place or what might have been said, having access to an archive of business information provides clear evidence of what goes on across an organization. Consider the following, proactive use cases for an archiving solution:

- **Better understanding customers**
  Customer relationship management (CRM) systems are vital for salespeople and others within an organization, but they don't tell the entire story of how the sales and nurturing process work. For example, a CRM system won't

Having access to an archive of business information provides clear evidence of what goes on across an organization.

necessarily include information about when a prospect's inquiry came in through email or some other communication channel. They don't always provide the tools that can track the relationship between the length of time it takes to respond to a customer inquiry and renewal rates for those customers. Salespeople, accessing and analyzing data from the corporate archive, might find useful differences in their sales staff performance that wouldn't show up by analyzing just CRM data.

- **Detecting policy violations**
  Organizations generally have a large number of policies that focus on how communication tools are to be used, which ones are not to be used, and so forth. Many times these policies are driven by regulatory requirements, such as the GDPR. An archiving platform can determine if employees are acting in violation of these policies and can alert managers that corrective action needs to be taken, both to address employee behavior and to protect the organization from failing to satisfy its compliance obligations.

- **Finding insider threats**
  Supervisors who belittle their employees or complain about them are more likely to create an environment for data breaches among those whom they manage. Disgruntled employees are not happy employees and are more likely to steal data, start their own business out of their cubicle, send sensitive information to competitors, and the like. An analysis of archived content can help to highlight these managers' behavior and address their abuse before employees reach the point of harming the company.

- **Monitoring language**
  If employees use profanity or harassing language in the workplace, including their email and other communications, it can create problems. For example, in Meritor Savings Bank v. Vinson[2], a court found that "even a woman who was never herself the object of harassment might have a sexual harassment claim if she were forced to work in an atmosphere in which such harassment was pervasive." An archive can be a useful tool for managers to perform searches on their employees' communications to determine if they are using language in their written communications that could somehow become actionable.

- **Understanding employee sentiment**
  A robust archiving platform can help to identify problems so that violations of corporate policy, legal requirements, or industry best practice can be addressed proactively before they result in a more serious problem. For example, a line-ofbusiness or HR manager could review archived data to understand employee sentiment related to recent company actions or to determine if morale is poor.

---

2 https://www.oyez.org/cases/1985/84-1979

Best practice recommendations

Osterman Research recommends a number of steps that organizations should consider as they consider how to implement a robust archiving capability as part of an overall information governance program.

Establish an executive sponsor

The initial step is to establish an executive sponsor for an information governance program – in essence, a champion that will push hard for good information governance and will serve as something of an evangelist for the program. This person might be the chief legal counsel or a compliance officer, or it could be an IT manager or the CIO, but there needs to be an executive sponsor that will push hard for the program despite the inevitable obstacles that will come up. The survey we conducted for this white paper found that while only 66 percent of organizations currently have an information governance program, 73 percent have an executive sponsor for their current or upcoming information governance initiative.

Understand the drivers and requirements for information

It is essential for decision makers to understand why an information governance program is necessary in the first place and to determine what is going to be driving the program going forward. For example, an organization that has a substantial number of customers within the European Union will have as one of its key drivers the GDPR. Organizations that are involved in eDiscovery on a regular basis will have as one of its information governance drivers the need to make eDiscovery more costefficient and effective. Understanding the drivers will help to determine who might be the best candidate for executive sponsor of the program.

Understand why archiving is important

Many organizations don't have coordinated strategies for retaining business content. Some will leave retention up to individual users or departments, retain all data for a certain length of time, or will purge data stores on a regular basis without enough regard for retention requirements for specific types of information, and so on. Even so, archiving is still a critical best practice and organizations should be retaining their business records for the appropriate lengths of time as prescribed by statutory requirements, regulators' demands, court precedent and legal counsel's advice. Every organization across all industries should implement appropriate and granular retention schedules – and a robust archiving solution – to ensure they are maintaining the appropriate defensive and functional posture for good data management.

Evaluate more proactive archiving

Organizational decision makers should appreciate the wealth of information that their organization has available to them if they properly archive their data and implement the right analytics for understanding this data. Archived information can reveal insights about the organization that simply are not available from any other source. Further, since relatively few organizations today have taken a next-generation approach to archiving, there will be a period in which those early to do so will gain significant competitive advantage by understanding their business more thoroughly than their competitors understand theirs.

More decision makers and influencers over time are viewing the ability to extract insight and intelligence from archived data as an important or major reason for having their archiving platforms.

As noted in Figure 5, more decision makers and influencers over time are viewing the ability to extract insight and intelligence from archived data as an important or major reason for having their archiving platforms. Consequently, archiving solutions in the future will increasingly be called upon to have this capability, something that not all archiving platforms can do today.
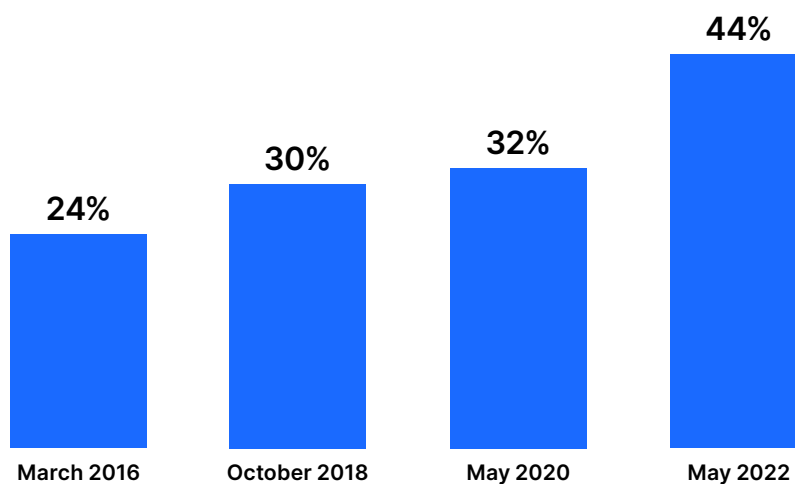

Figure 5 Importance of Extracting Insight and Intelligence from Archived Data
Percentage responding an "important" or "major" driver
Source: Osterman Research, Inc.

Implement the right archiving solution

Organizations must deploy the appropriate archiving solution. Among the key questions for an archiving solution are the following:

- Does the solution archive all of the electronic content that an organization will need to retain? For some organizations, this will be email only, while for others it will be a much broader set of content, such as text messages, social media posts, voicemails, content from collaboration systems like Slack and Microsoft Teams, and so forth.

- Does the solution archive content in context, not translating it into other, nonnative forms that strip the original content of its context and metadata?

- Is the solution scalable? Does it store and enable rapid searches of as much content as will need to be archived? It is important that archiving solutions should offer fast, reliable and unrestricted export of content. Many archiving solutions are slow to get data out and, in for some cloud-based solutions, will charge for data export.

- Does the archiving solution store data in a holistic way rather than in separate silos of content that must be pieced together by those extracting content from the archives?

- Where will the data be stored? Data residency – i.e., the location(s) in which archived data will be stored – is a key decision point because some jurisdictions require data to be stored only in certain geographic locations. Some customers are concerned that some nations' statutes are too far-reaching, and so do not want their data stored in those geographies.

It's important to keep in mind that not all archiving solutions are equally capable. With the move to Microsoft 365, for example, many organizations are relying on its native archiving technology, Exchange Online Archiving, to meet their archiving requirements. While that will work for some organizations, it's important for decision makers to perform robust due diligence to determine if Microsoft 365 – or whatever other platforms are used across the organization – will satisfy their current and future archiving requirements.

Summary

Archiving is an essential best practice that every organization should pursue to protect against the inability to find critical data during litigation, regulatory audits, or simply when employees need older data. Moreover, while archived data can be used for these types of defensive purposes, the same data set can also be used in a more proactive way to better understand various processes and practices within an organization. The benefits of using archived data in this way can lead to competitive advantage, risk avoidance, improved policy compliance, and a better understanding of how an organization operates.

Sponsored by 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.

opentext™