ELECTRONIC FILE CLEANUP PREPARE NOW OR PAY LATER!

A Hidden Cost is Revealed

Despite a decrease in the cost of electronic data storage, other costs associated with electronic information continue to rise. In most organizations, unstructured data is not effectively managed. From years of obsolete and duplicative documents on shared drives and in email to legacy backups, the unnecessary cost of storage, the threat of data breach, and the risk of e-Discovery are real. As piles of electronically stored data continue to grow, compliance and security become increasingly difficult to manage. Consequently, information becomes increasingly difficult to effectively retrieve.

The complexity and volume of this task can initially seem overwhelming, but don't sweep it under the rug because your cleanup project will only get more costly and complex over time. And if you think cleanup is expensive and time-consuming, wait until you go through the e-Discovery process! In this article, we will show you how to tackle electronic records cleanup without breaking the budget or your colleagues.

The Best Offense is a Good (Disposition) Defense

Before you begin your cleanup efforts, you must have the legal foundation to do so. Ad hoc or random destruction of records/information based on whether your colleagues think they need it or not opens up your organization to unnecessary legal risk. If records subject to a future legal discovery are missing, your organization can be at risk for a finding of spoliation of evidence, which can be quite costly – both in terms of money and reputation.

To ensure the legal defensibility of your cleanup efforts, you must have the two foundational elements of every records and information management (RIM) program in place – the RIM Policy and Records Retention Schedule. Therefore, the first step in cleanup is to review your Policy and Schedule to ensure that they reflect your current organizational structure, regulatory environment, and industry best practices. Even better, make sure you have a robust training program in place to ensure your colleagues know how to stay in compliance. After all, your information management program is only as strong as your weakest (untrained) link.

Once you have the foundational elements in place to operate a legally defensible RIM program, you may now turn your attention to the organization of your electronic records. A key step in organizing enterprise information is the development of a functional taxonomy. A taxonomy is essentially the structure and language of your organization's information. A functional taxonomy is important for the usability of your Records Retention Schedule and is leveraged as you develop file plans for your electronic information. No matter how sophisticated your system's search capability, defining a hierarchical structure by which your users can store and retrieve information improves your results and bolsters organizational productivity.

There also needs to be enterprise level guidance regarding file plans and naming standards. The department coordinators understand their business needs, but don't expect them to understand the complexities of applying compliance requirements, implementing a sustainable lifecycle management plan, or mitigating the risks associated with immature rights management plans. Additionally, in today's cyber environment, it is critically important to make sure you understand the security requirements associated with the information in your organization's electronic file plans.

Before building your file plan, gather everything known about your organization's technology strategy for managing unstructured information day-forward. You want to understand the strategy for all components of your unstructured data (e.g., documents, email, and web content). What software is part of your 5-year IT strategy and what software is targeted for upgrade or conversion? If there is a preferred content management tool, how do the retention policies integrate with the tool, and what implications does it have on your file plan?

Remember, records life cycle management is important whether you are implementing a single purpose repository (like clinical studies or web content) or a large multi-function, enterprise-wide solution.

Finally, if you are moving the active content to a new Content Services Platform (CSP) or Electronic Document and Records Management (EDRM) repository, or collaboration space, there are additional considerations such as metadata, site governance, and Search Engine Optimization (SEO).

Electronic File Cleanup Strategies

Once you have addressed the policy, records retention schedule, taxonomy, file plan and technology strategy, you can begin managing the cleanup of electronic information. There are several approaches that you can take:

1. I CAN DO IT MYSELF:
Certainly, many organizations simply delegate the task to the departments who own the information. Unfortunately, this is often done without completing the necessary preparation steps, discussed above. Without a solid foundation and go-forward plan, departmental cleanup efforts may result in a reduction of the well-organized information that is eligible for destruction, but typically an avoidance of the older or disorganized content that no one is sure about. Without a file analysis tool, this is a time-consuming process and business priorities justifiably win out at the expense of the cleanup effort.

2. THE IT APPROACH:
The IT approach simply reviews your content based on the created and last modified dates and targets a list of older content for destruction. There are obvious compliance issues with this approach. However, you can offset these issues with an approval process. Be warned that implementing approval processes (which are typically manual) is usually labor intensive for the business units and therefore often ignored. Conversely, approval completed by IT and/or Legal without full disclosure to the owners of the information can interfere with business processes, too.

3. DATA DISCOVERY AND RECORDS MANAGEMENT TOOLS:
There are many software tools that can help you analyze your shared drives and other unstructured repositories and make the process of cleaning up your electronic information easier to manage and legally defensible. The tools can be categorized as:

    a. DATA DISCOVERY or SEARCH:
    There are many sophisticated search engines out there that can help you analyze your shared drives, from standalone search tools or storage management tools to

and a long list of eDiscovery vendors. Commentary to Sedona Principle 11 notes: "The selective use of keyword searches can be a reasonable approach when dealing with large amounts of electronic data… This exploits a unique feature of electronic information - the ability to conduct fast, iterative searches for the presence of patterns of words and concepts in large document populations." These tools are usually leveraged with the intent of a more thoughtful consideration of which content should be deleted or preserved, but are only effective if testing, sampling, and iterative feedback are employed to move a low (e.g., 20% effective) "Go Fish" approach to something that can be over 80% effective. But again, compliance issues with this approach are often offset with an approval process that is typically manual - and labor intensive for the business units or completed by IT and Legal without full disclosure to the owners of the information.

b. **RECORDS MANAGEMENT SYSTEMS:**
Electronic Document and Records Management (EDRM) systems within Content Services Platforms (CSPs) have the tools to apply retention policy to your unstructured records. For the content in your CSP product, annual cleanup activities can be automated, compliant, and auditable. But few people want to spend the time and energy loading their old content into an CSP just to document the destruction process. Some of the vendors have data discovery modules that can search, auto-classify, and help organize content to eliminate the low hanging fruit before content is uploaded into a document and records management repository.

c. **IN-PLACE FILE ANALYSIS & CLEANUP:**
There are also tools that offer in-place cleanup of unstructured repositories (e.g., Rational Enterprise and Active Navigation), and the rebranded Microsoft Purview (formerly Security & Compliance Center), all of which can perform discovery, autoclassification and records lifecycle management functions. These tools leverage machine learning against user fed datasets to automate data classification, discovery, and lifecycle management. Be cautious, however, as these tools are not a magic bullet; you will need to invest the time to train these tools to achieve the appropriate level of accuracy. This can be time-consuming, and the process must be repeated as the data landscape changes within your organization.

The right tool for you will depend on your volume of information (terabytes or petabytes), your long-term IT strategy, your budget, and your need for compliant and auditable records cleanup. Which approach and potentially which software is best for you to use when cleaning up your information stockpile requires an understanding of your organization's volume of stale information, the technology strategy and budget, and the amount of labor that will be available for implementation. Then make sure you have a management-backed strategy, a retention policy, updated retention schedule, mature taxonomy and file plan, and a knowledgeable information governance team.

Need help?

Cadence Group has been providing information governance and records retention consulting services to public and private sector organizations of all sizes for over 30 years. Our certified consultants can help bring your organization into compliance by providing:

- A full information governance program assessment
- Records retention scheduling, including legal research
- Records inventorying and information mapping
- Physical and electronic records cleanup

Click here to set up a free consultation.