

ELECTRONIC RECORDS CLEANUP

AIIM 2013 National Conference - Session 03-21-2013

Tina Teree Baker, President/CEO Cadence Group

Despite a decrease in the cost of electronic data storage, other costs associated with electronic information continue to rise. In most organizations – from global to mid-sized, corporate to federal – unstructured data is not effectively managed. From years of obsolete and duplicative documents on shared drives and in email - to vintage 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. And information becomes more and more difficult to effectively retrieve.

The complexity – and volume – of this task can initially seem overwhelming. Don't sweep it under the rug. If you think cleanup is expensive, wait until you go through the e-discovery process. Here's the first thing you need to do:

Review your organization's Records Retention Schedule to make sure the categories address the information you store electronically (as well as hardcopy) and that it is compliant with current laws and regulations.

After having clarified and leveraged a functional taxonomy as you reviewed (or created) your organization's records retention schedule, you are positioned to develop a file plan. A functional taxonomy is important for the usability of your RRS and is leveraged again 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.

There needs to be corporate level guidance regarding File Plans and naming standards. The department coordinators understand their business needs, 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. Make sure you understand the security requirements associated with the information in your organization's electronic File Plans.

Now, 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 document management tool, how do the retention policies integrate with the tool, and what implications does it have on your file plan?

Remember, 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, I want to note that if you are moving the active content to a new Enterprise Content Management (ECM) or Electronic Document and Records Management (EDRM) repository, or Collaboration space, there are additional considerations like metadata, site governance, and Search Engine Optimization (SEO).

Now, we are ready to talk about 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. Without a good plan, departmental cleanup efforts may result in a reduction in 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 tool, this is a time consuming process and business priorities justifiably win out.
2. **THE IT APPROACH:** what I like to call the IT approach has been around since my old days of designing systems and managing data on mainframe computers. This approach simply reviews your content based on the created and last modified dates and targets a list of older content for destruction. The obvious compliance issues with this approach are sometimes 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.
3. **DATA DISCOVERY AND RECORDS MANAGEMENT TOOLS:** There are many software tools that can help you analyze your shared drives and make the process of cleaning up your electronic information easier to manage and legally defensible. One might choose to categorize the tools as:
 - a. **DATA DISCOVERY or SEARCH:** There are many sophisticated search engines out there that can help you analyze your shared drives, from stand-alone search tools or storage management tools to and a long list of e-discovery 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 (let’s say 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 ECM suites have the tools to apply retention policy to your unstructured records. For the content in your ECM 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 ECM tool just to document the destruction process. Some of the vendors (e.g., IBM/StorageIQ and EMC/Kaseon) 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 CLEANUP:** There are also tools that offer in-place cleanup of shared drives (e.g., ePurge, Rational Retention, and Active Navigation). The right tool for you will depend on your volume of 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, mature File Plan, and a knowledgeable project manager.

Whatever the decision might be, it needs to be made soon.