CMS WiRE

Here's How You Clean Up Network Shares

By Joe Shepley | *May 31, 2018*

Follow 1,360 followers

How AI Technology Can Address M

PHOTO: NUNO SILVA

Despite the increasing capabilities of information management technology and despite how long organizations have been trying to implement better information management, the most frequent question I get from clients is "How do we clean up network shares?"

And despite seeming like something that should be easy to do, it is anything but that. Which isn't to say it's complex — it isn't. But like diet and exercise, which are simple in concept but arduous in practice, cleaning up your network shares requires a simple approach with a ton of elbow grease.

In this post, I want to lay out an approach for cleaning up network shares that has been successful at organizations of varying sizes across a range of industries.

A 3 Step Approach to File Share Cleanup

The first step is to acquire file analytics software to help you evaluate the documents on your network shares. Even a small organization will have a few terabytes of content on shares, which translates into millions of documents — far too many to evaluate manually.

The good news is a range of file analytics tools are out there, from the dead simple (like a Tree Size Professional, which costs under $100) to the extremely sophisticated (e.g., Active Navigation, Nuix, ZL Technologies, StoredIQ, Concept Searching, etc.). So, depending on your budget and maturity level, you'll find technology to assist you in your network share cleanup efforts.

Regardless of this diversity in file analytics tools, one cleanup approach is successful more often than not:

> **Step 1**: **Address ROT —** Finding and addressing redundant, obsolete, and transitory documents is the most straightforward step and requires the least horsepower from a file analytics tool.
>
> **Step 2: Address Sensitive Data** — Finding and addressing documents containing sensitive data such a s PHI, PII, PCI, IP, legal hold data, etc. requires more sophistication from file analytics tools.
>
> **Step 3: Address Records** — Finding and addressing records requires the most sophistication from file analytics tools.

Let's walk through each of these in turn to see how they might be valuable for your organization's network share cleanup efforts.

Related Article: Fixing Your Shared Drive Problem

Step 1: Address ROT

In order to address ROT, a file analytics tool only needs to interrogate file properties (so called wrapper metadata), i.e., file path, file name, file type, file size, date created, date last accessed and date modified. It does not need to crack open the file and interrogate the contents. For this reason, scans are faster (with speeds approaching 1,000,000

documents an hour). In addition, anywhere from 30 to 80 percent of network shares at an organization will be ROT, so performing this analysis first takes a significant amount of documents off the table for analysis in later, more complex steps.

Once you find the ROT, you need to do something with it, which can be as simple as quarantining it (whether in a dedicated repository or by leaving in place or removing access) or as final as deleting it. What an organization does in any given case depends on organizational culture, the policy and procedure landscape, etc., but finding ROT first gives you the ammunition to take the first step — and paves the way for addressing more valuable and risky network share content.

Step 2: Address Sensitive Data

Once you've addressed ROT, you've likely reduced the volume of your network shares by 30 to 80 percent, which will make finding and addressing sensitive data much less resource- and time-intensive.

Typically, sensitive data at any organization would include:

PHI – protected health information

PII – personally identifiable information

PCI – payment card information

Depending on the organization, sensitive data could also include:

Board materials

Intellectual property

Financial analysis

Mergers and acquisitions

Although the approach is straightforward (i.e., find sensitive data and secure it), the devil is in the details — and there are far too many of them to address exhaustively here. But suffice it to say, you'll use a combination of out of the box and customized rules for identifying sensitive data using pattern matching (e.g., ###-##-#### as a pattern suggests a social security number may be present). It's quite time intensive, even with out of the box rules, because the goal is to find the sweet spot between too broad and too narrow, and that can take many iterations.

Related Article: Do You Know What's In Your File Shares?

Step 3: Address Records

With ROT and sensitive data addressed, you can move on to addressing records … or not. Once you reach this point, you may find only 5 to 20 percent of your network shares remain, so it might not be worth the effort to go further. But if you do, know there's no easy way to make progress: it'll be a long, hard slog, one department (sometimes one record series) at a time.

Some thoughts on how to make this slog a little less difficult:

Start by rationalizing your records retention plan, the goal being to get to as few categories as possible (e.g., if a department has three, five and seven year retention periods for documents, make them all seven and you can classify their documents all as one thing … and perform disposition on them much more easily)

Find departments that organize their network share by record types. If you can find folders that have a single record type in them, you can classify and disposition them more easily (e.g., finance, HR, compliance, M&A, etc.)

Find record types that can be easily identified using pattern matching. Then you can create rules à la sensitive data to find records (e.g., contracts, POs, invoices, etc.)

An Easy Concept That's Hard to Execute

As I said at the opening, cleaning up network shares is conceptually straightforward but difficult in practice. Hopefully the approach I've laid out gives you the clarity on how to proceed and convinces you that, given the right technology, vision and elbow grease, you can make meaningful progress on cleaning up your network shares.

About the Author

Joe Shepley is a strategy consulting professional living and working in Chicago. In his current position as Vice President and Practice Leader at Doculabs he focuses on helping organizations improve how they manage information using technology and processes.

Jun 26 [Nuxeo Webinar] Ask the Analyst: 6 Important DAM Capabilities

Jun 26 [CMSWire Webinar] How to Make Enterprise Budgeting Easier

Jun 27 Modeling, Measuring and Optimizing the Customer Journey #DXChat