A White Paper sponsored by Active Navigation

Kill Your File Shares

A Roadmap to Reducing Your Biggest Information Risk

By Russell Stalters, CEO Clear Path Solutions Inc.

There are many good reasons to kill your file shares, so what holds organizations back from tackling this challenge? When referring to corporate file shares, I am including shared network drives, personal file shares, cloud storage and file sharing applications.

Think about it. As a business leader, if you were walking around the offices of your company and saw stacks of papers, documents, video DVDs, and music CDs piled high on desks and laying on the floor around the desks, you would be appalled. You would immediately tell your employees to clean things up, throw out the trash, and organize the materials that they needed to do their job to support the company's mission! It's a fire hazard and something important could get lost. Right?!

Well, that is what most companies' corporate file shares look like. They are filled with electronic trash (eTrash) which can include non-business related copyrighted music files, empty files, old backup files, old application setup files, games, and employee personal photographs. They also contain redundant, obsolete, trivial (ROT) data, documents, and files. Our colleagues and employees are treating these corporate file shares like dumping grounds.

One reason organizations don't clean up this mess in corporate file shares is the "out of sight, out of mind" problem. Most of the eTrash and ROT is hidden away from plain sight. Some file shares are only known to a small group of people creating the notion of Dark data. This data is typically unstructured, untagged and untapped data that is found in corporate repositories and is mostly neglected by business and IT administrators in terms of its value.

Over the years, **Russell Stalters**, CEO of Clear Path Solutions Inc. and author of gettinginformationdone.com, has established a reputation as a subject matter expert for information governance, information and data management, SharePoint and Office 365 based solutions. Most notably until 2016, as Director Information & Data Management for BP's Gulf Coast Restoration Organization, he was responsible for information architecture, data management strategy and geographic information services. In that role he served as the Chief Architect and data strategist for all business applications and data management solutions created to manage data collected during the response and restoration efforts from the Deepwater Horizon Incident and Oil Spill.

He also helps organizations innovate, transform, and maximize the effectiveness of individuals by helping them improve their ability to lead, work together, select, and develop their people. He has honed these skills as an Executive Director with the John Maxwell Team and is personally mentored by John C. Maxwell.

A regular industry keynote speaker, noted author, and thought leader, he presents at national and international industry conferences. In 2014 he was inducted into the Association of Imaging and Information Management (AIIM.org) Company of Fellows to recognize him as an expert in the field of information management. Russell has a master's degree in Computer Science from the University of Florida and is a retired Naval Officer and Naval Aviator.

Contents

The Case for Killing Your Corporate Files Shares

If you don't know what is stored in your corporate file shares, how can you begin to understand the potential risks that lurk inside them? It is time to take bold action and "kill your file shares." I know, it sounds drastic. By "kill your file shares" I mean: kill off the practice of storing data in random file shares scattered throughout the company and begin formulating an information governance initiative and remediation plan today. Doing so will not only immediately reduce your business risk, it will increase employee productivity and your company's speed and agility by leveraging corporate information assets more effectively.

Last year, I published the white paper titled: ***Not if, But When You Get Hacked: Measuring and Proactively Managing Information Risk***. I discussed how, without a robust Information Governance Program and an effective way to measure its effectiveness, organizations can be exposed to significant information risks which contribute to overall enterprise risk. The information risks and the impact they can have on an organization included:

- Protection of Intellectual Property and Sensitive Corporate Data
- Legal Hold and eDiscovery Processes
- Information Quality
- Data Breaches and Privacy Compliance
- Proliferation of Information Systems and Repositories

The mindset I proposed senior executives and business leaders adopt is:

"when we get hacked" not "if we get hacked"

With some of the most recent high-profile data breaches including Equifax, Deloitte, and Verizon affecting over 100 million people now is the time to take bold action and significantly reduce the information risk lurking within most companies. But, where to start!

I recommend starting with one of the biggest contributors to enterprise risk from information. That is toxic ROT, eTrash, and hidden sensitive data. And where do the greatest volumes of risk exist?

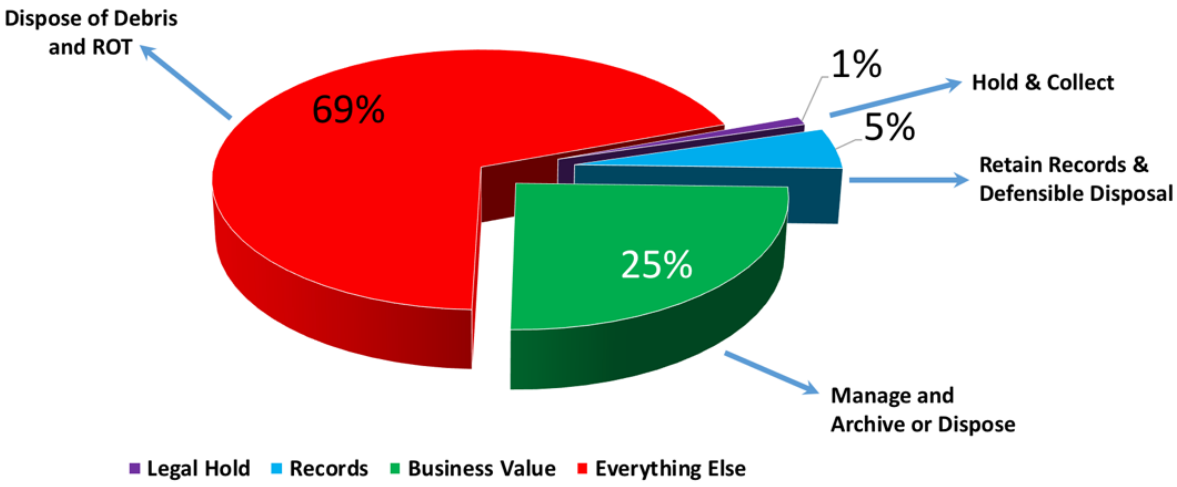
Corporate File Shares

Much of the data residing in these corporate files shares is considered Dark Data. An Osterman Research White Paper concluded the vast majority of data accumulating within the typical enterprise is unstructured and usually **controlled and "managed" by individual employees**. The paper suggested this data be considered **"Dark Data"** because it is invisible and not easily accessible by the company[1].

How can senior management and the Board know the magnitude of problem and the impact on the organization's enterprise risk profile? Show them and do something to reduce that risk quickly.

How Big is The Problem?

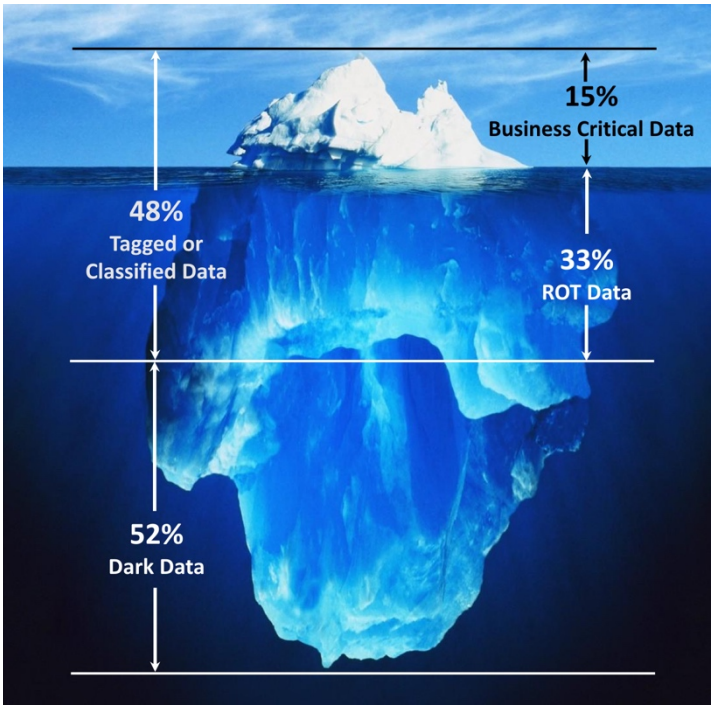
This enormous problem is pervasive throughout the enterprise as demonstrated in this chart which shows that almost 70% of the information across enterprise repositories are eTrash and ROT[2]. How much of this information is sensitive data?



These are the categories of highly sensitive data subject to regulation that constitute the majority of the costs associated with responding to and remediating a data beach:

- Personally identifiable information (PII)
- Sensitive personal information (SPI)
- Protected health information (PHI)
- Intellectual property (IP)

A study by Veritas Technologies surveyed over 2,500 respondents from 22 countries and found that only 15% of data was identified as vital to the on-going operational success of the business while a whopping 85% was classified as ROT and Dark Data. They defined Dark Data as data that may include business critical data as well as eTrash, or ROT, and may contain non-compliant data leading to significant and unseen business risks[3].



What's in Your Corporate File Shares?

The first question you need to answer: how much toxic ROT, eTrash, or Dark Data is being stored in corporate file shares? Another important question is how many of the file shares are in the enterprise that contain ROT and sensitive data?

To answer the first question, we need to have an accurate data map of the corporate file shares and the volumes of the data stored on those drives, as well as personal file sharing and cloud based storage accounts. Recent studies indicate companies continue to liberally use these drives. A Benchmark by the Information Governance Initiative found that corporate file shares are the most common repository for the storage of information we know must be protected and accessed for at least ten years. Sixty eight percent of the respondents, the top response, identified them as a storage location[4].

A case study of Pandora Media illustrates this. The Pandora team determined that 60 percent of the content on Pandora Media's corporate file shares was redundant, obsolete, or trivial. As a result of eliminating the ROT, the company immediately reclaimed 18 terabytes of expensive data storage, at the same time reducing the risk associated with holding that data long-term[5].

Another case study of Les Schwab, one of the nation's oldest and largest retailers and distributors of tires and related automobile parts, shows how quickly corporate file shares can get out of control. They recently passed the 5 terabyte (TB) mark in its shared drive environment alone. Department shared drives (i.e., those shared by members of a group as opposed to network file shares assigned to individuals) had over 500 top-level folders, and thousands of additional subfolders. *Nearly half* of that information had not been accessed within *the past five years*. The team quickly identified 1.7 million files that were ROT and could be deleted[6].

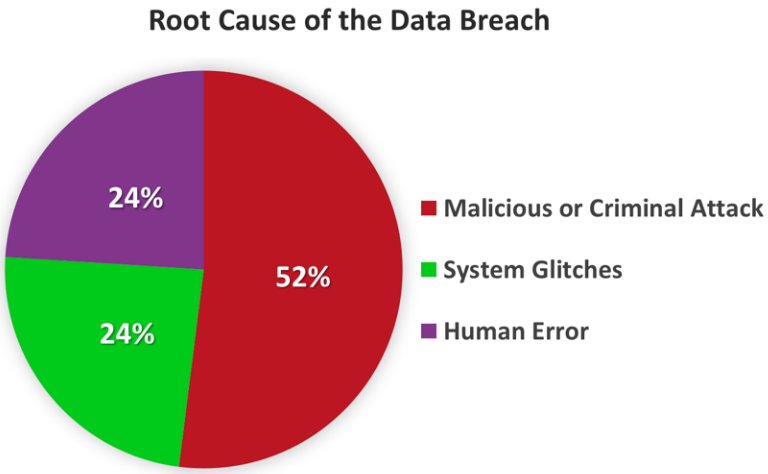
Understanding the Magnitude of the Risk

Many companies have multiple file shares which contain hundreds of terabytes of uncategorized and typically unknown content. eDiscovery Directors and Information Governance Directors mention this as one of the principal risks they want to get after. They recognize the fact that much of the content contained in these repositories poses a significant litigation, regulatory, and compliance risk.

If we can convince the C-Suite and senior business leaders to adopt the mindset of "**when we get hacked" not "if we get hacked"** we can show them the value of remediating this risk. Also, these risks are more than just getting hacked or a data breach. For example, a lawsuit or governmental investigation can result in unnecessary excessive costs because Legal had to collect and review data that should have been responsibly deleted years ago.

The costs are real when a data breach does occur. The most recent Ponemon Institute "Cost of Data Breach Study: United States" found the average cost for each lost or stolen record containing sensitive and confidential information **is $225** and the average **total cost for a data breach is $7.35 million**[7].

But even more staggering is that a breach, which includes lost or compromised sensitive records, does not have to be the result of a malicious or criminal attack! While the study found that *52 percent of incidents involved a malicious or criminal attack, 24 percent involved negligent employees and another 24 percent involved system glitches that include both IT and business process failures.*



Source: 2017 Cost of Data Breach Study: United States[7]

Additional Components of Risk

The following are some additional resources to help understand and explain to stakeholders the magnitude of the risk:

1. Risks related to health and personal health information subject HIPAA: Penalties for noncompliance are based on the level of negligence and can range from $112 to $55,910 per violation (or per record), with a maximum penalty of $1.68 million per year for violations of an identical provision[8].
2. Each state (except Alabama and South Dakota) has laws and regulations to cover the unlawful and unauthorized release or loss of personal information that compromises the security, confidentiality, or integrity of personal information. The penalties vary from state to state. For example, the Attorney General of Massachusetts can seek a penalty of $5,000 for each violation, and reasonable costs and attorney's fees[9].

3. New York State companies operating under or required to operate under a license, registration, charter, certificate, permit, etc., under the State's Banking, Insurance, or Financial Services laws are required to maintain a cybersecurity program[10]. Some of the company's responsibilities include:
    a. Have a cybersecurity policy based on the company's risk assessment
    b. Designate a Chief Information Security Officer (CISO)
    c. Conduct a risk assessment which addresses the following areas (to name a few): data governance and classification; asset inventory; customer data privacy; incident response
    d. Senior management must take this issue seriously and file an annual certification confirming compliance with the regulations
4. For companies who collect and process personal data from individuals residing in the European Union, the General Data Protection Regulation (GDPR) represents a significant potential risk if the company has a breach. Under GDPR organizations in breach of the regulation can be fined up to 4% of annual global turnover or €20 Million ($23.4 million at the time of this writing), whichever is greater[11].

Quantifying the Risks of Corporate File Shares

The is no perfect way to quantify the monetary value of the risk posed by corporate file shares. However, I propose an approach we can use to estimate the financial exposure based on the studies referenced in this paper. This is merely a starting point but will generate something that you can share with business leaders to highlight the potential risk.

In the Cost of Data Breach Study: United States *the number of breached records per incident this year ranged from 5,563 to 99,500 records. The average number of breached records was 28,512*[7].

Each company is different and the distribution of toxic ROT, eTrash and sensitive information across the landscape and will differ by industry, size, age of the company, and total volume and number of records stored on corporate file shares. What we can say is that based on the results from the study the potential exposure in USD (using the average cost of $225 per record) could range anywhere from $1,251,675 to $6,415,200 using the low number and the average number of records breached[7].

Potential Exposure in Dollars:

$1,251,675 to $6,415,200

Non-Malicious/Non-Criminal Breach Could Cost Organizations on Average Approximately $1,125,000

An IT administrator inadvertently and unknowingly changes the permissions on a corporate file share so that anyone authenticated on the network has access to all the files. One file is an Excel spreadsheet that is an extract from the HR system with over 5,000 employees PII including SSNs, salary information, and dependents information. There is no way to know how long this breach has been there and how many have accessed this file. The organization now has to assume a full breach and contact the employees in the spreadsheet, many who have left the organization and provide credit/identity protection. Based on the study above this can cost $1,125,000 just for a poorly executed IT process and the breach of one

A Roadmap for Reducing Your Biggest Information Risk

Having made the case that corporate file shares pose one of the biggest information risks in the company, now let's do something about it. The following is a roadmap for attacking this problem head on and fast. Active Navigation calls this type of initiative; "Kill Your File Shares".

The approach I propose can be tailored to your organization and may vary based on the size of the organization (number of employees), the geographical footprint, and the industry. Smaller organization may want to apply this roadmap enterprise-wide, while larger organizations or ones which operate globally may want to start by applying this roadmap to a department or portion of the organization first.

Education and Executive Buy-in

The first step is to help company executives including the C-Suite, board members, Risk Committee, and other key stakeholders understand the magnitude of the risk. Ways to do this include:

1. Share my first white paper; ***Not if, But When You Get Hacked: Measuring and Proactively Managing Information Risk*** and this white paper. **http://bit.ly/WhenHacked**
2. Share other instructive resources to amplify the potential risks and impact to the company, including recent data breaches and associated costs as well as the resources listed in this paper. Most will not read them completely, so I recommend creating a one-page Executive Summary which summarizes the listed references.
3. Create a presentation highlighting the risks specifically related to your company and provide some indicative risk exposure amounts in dollars. This can be based on the number of files in the company or selected department file shares using the data provided earlier in this paper. IT should be able to provide this information. We already know these figures will not be completely accurate, but it's a place to start and will spark discussions with executives. For example, an Energy or Utility company should be concerned about customer data (PII and PCI), health and safety data for safe operations, and environmental data to maintain a "license to operate" at the federal and state level. The presentation should highlight the risks associated with a breach or poor or missing data, and the potential business impacts to the company.
4. Find executive allies or champions within the organization that understand the risk and are willing to help you take action. This might be a department head who is willing to sponsor a pilot to remediate the file shares for their department. Other candidates include the General Counsel, Chief Information Security Offices, CIO, and Chief Risk Officer.

Create a 'Kill Our File Shares' Proposal

Create a proposal for a fast, small "Kill Our File Shares" project to significantly reduce ROT and eTrash along with sensitive information and the associated risks. Don't try to "boil the ocean" or "eat the elephant" by trying to solve this enterprise-wide all at once. Create the project proposal with that ally or champion you enlisted in the step above. Present the ROI in terms of the reduced impact of a data breach along with a side benefit of reducing storage. Use the total cost of managing a terabyte of file share capacity at $3,000 per year, every year[5]. This immediate economic benefit can also be part of the ROI.

Access the Data

In my last white paper, I made the case for creating a comprehensive data map with the following benefits:

- Profile the data estate to understand what you hold and design and act on policies driven by what data you actually have
- Locate, remediate, and adequately protect files containing sensitive information
- Identify the millions of ROT files, which once removed defensibly improve search quality and reduce risk
- Design information protection strategies based on the business value or associated risk of the information
- Help the Chief Information Security Officer (CISO) prioritize investments for protecting sensitive information

Creating the data map can be automated using a world class file analysis tool like **Active Navigation** to scan the files shares and help expose that Dark Data that lurks within them. In this case remediating file shares, whether enterprise-wide or for a particular department or category of data, will help you get started on building that data map. This will be an opportunity to enlist the help of the head of eDiscovery to help reinforce the benefit of this project for them.

Scan and Remediate the Data

This is where the magic happens. Most corporate file shares have many terabytes of content comprising millions of files. It is IMPOSSIBLE and cost prohibitive to try and use manual methods to review all this data. This is where file analysis tools like **Active Navigation** should be used to scan and remediate the data. Some of the benefits of using a solution like this include:

- Smart conditional file content analysis
- Rules to identify and classify leaked PII, credit card, protectively marked and other types of sensitive data
- Ability to customize rules to meet specific organization business needs
- Act in place to defensibly delete or quarantine unwanted data
- Notify, review and remediate sensitive data

> *"File analysis enables storage managers, legal and security professionals, and business analysts to understand and manage unstructured data stores to reduce costs and risk, increase efficiency of business-critical data, and make better information management decisions for unstructured data."* ~ Gartner[12]

This single action will expose the dark data and provide insight into what is hidden in those file shares. By removing the toxic ROT and sensitive data we can reduce the "risk surface area" that could be breached.

Once the data has been reviewed and remediated we don't want to stop there. You should leverage the file analysis solution on a regular (weekly/monthly) basis to proactively monitor the corporate file shares to alert and remediate if eTrash, ROT, or sensitive data are stored onto the clean file share again.

Training and Reeducation

Now that the file shares are cleaned up, we need to prevent them from becoming contaminated with more ROT, eTrash and sensitive data. Employees who use these drives need to be educated on the new policy discussed above and a new culture of information management excellence created. A great resource for the change that is needed is the book by Chip and Dan Heath – *Switch: How to Change Things When Change Is Hard.* There are other resources available regarding creating and installing a culture of information management excellence which I can share later.

Conclusion

By reducing toxic ROT, eTrash, and eliminating sensitive data in corporate file shares, organizations can realize real benefits and reduce the risk associated with this pernicious problem. Immediate benefits include:

- Increased business speed and agility: the time it takes to locate a file among multiple versions stored across multiple drives impedes the pace of business and makes it difficult to act quickly when opportunities arise.
- Improvement of information quality by removing redundant information made up of previous versions of documents, numerous drafts, most of which have no value once the final version is published.
- Reduction of the eDiscovery footprint which reduces the costs associated with collection, review and production.
- Potential for decommissioning of obsolete or underutilized storage and information systems. Annual cost savings for managing a terabyte of file share capacity at $3,000 per year, every year, and $40K per IT application[1] is conservative for most industries.

The big prize here is the reduction of one of the biggest contributors to information risk within the company. Additionally, by selecting corporate file shares which is a very discreet and easy to access repository, organizations can get results fast – in as little as two weeks!

Leveraging file analysis adds precision to understanding what is lurking in those corporate files shares and remediating the toxic ROT, eTrash and misplaced sensitive data.

Solutions like Active Navigation can help reduce your biggest information risk.

About Active Navigation

Active Navigation is a recognized industry leader providing unique file analysis software for the discovery, transformation and ongoing control of unstructured data wherever it lies in the enterprise.

Its products play a fundamental role in any information governance strategy enabling cost, risk and efficiency savings through information audit, clean up and defensible deletion, intelligent file migration, records capture, eDiscovery collection and ongoing policy monitoring. Its globally proven methodology is driven by 10s of thousands of hours of practical experience and empowers information, IT and legal professionals to truly understand and take control of their unstructured data.

Active Navigation has a history of early innovation and leadership in file analysis and information governance, achieving Gartner Cool Vendor status and various Microsoft platform certifications as well as being a founding member of the Information Governance Initiative.

www.activenavigation.com

Endnotes

1. “Best Practices for eDiscovery and Regulatory Compliance in Office 365®” An Osterman Research White Paper, Published September 2016
2. “Information Lifecycle Governance Leader Reference Guide” by Compliance, Governance and Oversight Council (CGOC), 2014.
3. “The Databerg Report: See What Others Don’t”, Veritas Technologies LLC, 2016
4. Information Governance Initiative, “The Governance of Long-Term Digital Information: An IGI 2016 Benchmark” (Information Governance Initiative LLC, May 2016)
5. Information Governance Initiative, “IG Snapshot: Pandora Media – How Pandora Tuned into Information Governance Control of Its Most Sensitive and Valuable Information Assets” (Information Governance Initiative LLC, August 2017)
6. Information Governance Initiative, “IG Snapshot: Les Schwab – How a Leading Retailer Got on the Road to IG” (Information Governance Initiative LLC, September 2017)
7. “2017 Cost of Data Breach Study: United States” conducted by Ponemon Institute LLC, June 2017
8. United States Code of Federal Regulations, 45 CFR 102
9. “State Data Breach Law Summary” BakerHostetler’s, July 2017
10. New York State Department of Financial Services 23 NYCRR 500: Cybersecurity Requirements for Financial Services Companies
11. The EU General Data Protection Regulation (GDPR), European Union, April 2016
12. “Gartner Market Guide for File Analysis Software”, Gartner, Inc. August 2015

