Building the Case for Defensible Deletion

Deleting unneeded data allows organizations to reduce storage management costs, speed up discovery and investigations, switch off obsolete storage systems and tame the Big Data beast.

For many years, organizations have followed the advice of their consultants or lawyers and kept all their data, often beyond the mandated retention period. But IT, compliance and records management staff recognize this is no longer an option, as data volumes grow to hundreds of terabytes – or even petabytes.

The financial and time costs of maintaining storage for so much data are becoming prohibitive. If they are not so already, they will be within two years. The volumes of data are too vast to apply records management and retention policies effectively. In addition, much of this data is unknown, posing significant business risks and adding to the time and expense of discovery or investigation exercises.

Defensible deletion is a repeatable and accountable process that allows organizations to identify, categorize and manage all their data across multiple geographical locations, applications and storage and archive systems. With this knowledge, an organization can allow data to expire if it has no business value or legal hold requirements.

Disposing of unneeded data allows organizations to reduce storage management costs, speed up eDiscovery and investigations, switch off obsolete storage systems and tame the Big Data beast. Defensible deletion is part of a broader process of storage optimization, which may also include remediating archives and migrating data to lower-cost platforms such as cloud storage.
THE CASE FOR DEFENSIBLE DELETION

The case for defensible deletion is easy to make. Compliance and records managers recognize the need to reduce data volumes so they can apply retention and records management policies to the valuable data. IT managers and C-level executives understand their organizations are paying a lot of money to store data that is of no value. They know instinctively that deleting data they no longer need will reduce costs. It might also make it easier for their organization to react to real business needs and for its staff to find information when they need it.

Despite this logical understanding, most companies are still paralyzed inside an ever-growing mountain of duplicated, outdated or irrelevant documents and emails. The main obstacle to pressing ‘delete’ is the perception that an organization cannot accurately discern valuable or actionable data from space filler. However, technology is now available that scales to the massive volumes of data that large organizations store. This technology empowers organizations to make defensible and cost-effective decisions to delete or retain data. Nuix’s defensible deletion solution applies automated techniques to classify data, with manual review for questionable cases, so organizations can:

- Identify redundant, obsolete and trivial (ROT) data
- Find, retain and profit from valuable data
- Highlight data that may present a business risk.

With this knowledge, organizations can create a set of deletion rules and processes to which all key stakeholders have agreed, allowing them to delete the data they don’t need, while retaining and proactively addressing the information that contains business value or risk.

WHY DO COMPANIES NEED A DEFENSIBLE DELETION PROGRAM?

Most companies that implemented email archiving systems believed this acquisition would solve their information governance problems. However, most organizations have not achieved the information governance benefits they would have liked because:

- They applied broad, unqualified retention periods
- They needed to classify their data as it was created, not after the fact
- Their archive indexing could not identify risky documents (such as those that contained personally identifying information or customers’ financial details) or perform complicated searches that identified similar or near-duplicate documents
- They usually had no way of implementing these requirements without expensive software upgrades or major services engagements.

This is why most companies find themselves with masses of email and data, much of which they know is ROT.

ASSET OR LIABILITY?

Is unstructured data an asset or a liability? In reality, it’s both. Or at least, some unstructured data is an asset and some is a liability. By eliminating the liabilities, organizations can then set about gaining value from their unstructured data assets.
THE BUSINESS CASE FOR DEFENSIBLE DELETION

Organizations can quickly build a solid business case for defensible deletion. They can easily demonstrate that defensible deletion can:

- Prolong the life expectancy of existing storage infrastructure
- Avoid or defer storage refresh and upgrade migration costs
- Cut the costs of managing storage.

The report, Defensible Deletion: Quantifying the Benefits by ESG analyst Katey Wood, identifies five-year cost reductions of up to $1 million for small organizations, $4.75 million for medium-sized organizations and nearly $20 million for very large enterprises. These figures are based on estimates that 50% of data in organizational stores is redundant, obsolete or trivial.

Defensible deletion can also make eDiscovery faster and cheaper. One organization that recently undertook a defensible deletion program reduced its eDiscovery costs by 95%.

Based on an estimate of $5,000 per gigabyte for eDiscovery costs and a legal or regulatory investigation addressing 1% of an organization’s data stores, ESG estimates savings of up to $2.5 million for medium-sized organizations, $12.5 million for large organizations and $50 million for very large enterprises, over a similar period.

Reducing these costs and eliminating storage management tasks can make an organization’s IT group much more nimble and responsive to business needs.

POTENTIAL STORAGE DELETION COST SAVINGS ESTIMATES OVER FIVE YEARS AT 50% REMEDIATION

<table>
<thead>
<tr>
<th></th>
<th>Low-end storage savings</th>
<th>Mid-tier storage savings</th>
<th>Tier-1 storage savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium enterprise (10–50 TB)</td>
<td>$30,000–$150,000</td>
<td>$60,000–$300,000</td>
<td>$100,000–$500,000</td>
</tr>
<tr>
<td>Large enterprise (100–250 TB)</td>
<td>$300,000–$750,000</td>
<td>$600,000–$1,500,000</td>
<td>$1,000,000–$2,500,000</td>
</tr>
<tr>
<td>Very large enterprise (1 PB+)</td>
<td>$3,750,000</td>
<td>$6,000,000</td>
<td>$10,000,000</td>
</tr>
</tbody>
</table>


POTENTIAL DISCOVERY COSTS OF 1% OF DATA

<table>
<thead>
<tr>
<th></th>
<th>Discovery costs for 1% of enterprise data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium enterprise (10–50 TB)</td>
<td>$500,000–$2,500,000</td>
</tr>
<tr>
<td>Large enterprise (100–250 TB)</td>
<td>$5,000,000–$12,500,000</td>
</tr>
<tr>
<td>Very large enterprise (1 PB+)</td>
<td>$50,000,000</td>
</tr>
</tbody>
</table>


ROOTING OUT THE ROT

In the average organization, 50–70% of the data they hold is ROT: redundant, obsolete or trivial.

- Redundant data includes multiple copies of the same document stored in file or email systems.
- Obsolete data is so old, it falls outside the organization's retention policy.
- The most trivial data is easy to identify – such as spam – but relevance is in the eye of the beholder.

Nuix defensible deletion solutions can identify, index, categorize and eliminate data that has no business or operational value. A safe, defensible data deletion policy means being able to understand the content of an organization’s key information stores, regardless of their location, file type or size.

Deleting unneeded data allows organizations to reduce storage management costs, speed up discovery and investigations plus switch off obsolete storage systems.
INFORMATION PAPER: BUILDING THE CASE FOR DEFENSIBLE DELETION

Organizations looking to put in place a defensible deletion program can apply the following principles to ensure they gain maximum value. They should:

• Identify what data they can delete and what data they must retain, and make content-driven decisions about the rest
• Put processes in place to make decisions about whether to retain or delete data that cannot immediately be identified as valuable or worthless
• Use sophisticated, graphical searches to make and validate more informed decisions
• Approach deletion projects one repository at a time, dealing with the low-hanging fruit first
• Bring in knowledgeable external counsel who can work with internal counsel through a defensible deletion process, and who will sign off at the end of the process that the data marked for deletion is acceptable to delete.

USE CASE
Financial institution saves millions of dollars with defensible deletion solution from Nuix

A multinational financial institution with servers in three continents was looking to extract, manage and migrate more than 300 terabytes of data from EMC EmailXtender and Symantec Enterprise Vault archives contained in EMC Centera storage. These archives contained more than 3 billion email messages, some of which the bank had to retain due to ongoing legal action.

The financial institution engaged Nuix in early 2012 to build a server and storage infrastructure capable of extracting and processing the data as quickly as possible. Nuix built the infrastructure and began extracting and processing the data using Nuix’s Enterprise Discovery software. The bank and Nuix completed this exercise within six weeks.

The company found more than half the data was duplicated and a large additional part of the data was had no legal risk or business value. Overall, more than two-thirds could be deleted before moving the remainder to a more cost-effective archive platform. By retiring its old email archives and minimizing the volume of data stored in the new archive, the bank expects to save millions of dollars in storage hardware, software and management costs.

TIPS FOR A SUCCESSFUL DEFENSIBLE DELETION PROGRAM

ABOUT NUIX

Nuix is a worldwide provider of information management technologies, including eDiscovery, electronic investigation and information governance software. Nuix’s powerful, intuitive technology makes it possible to search, investigate and actively manage unstructured data sets of any size or complexity. This enables organizations to respond quickly and effectively to litigation or regulatory action, mitigate risks, reduce costs and extract value. Nuix customers include the world’s leading advisory firms, litigation support providers, enterprises, government departments, law enforcement agencies and all of the world’s major corporate regulatory bodies.