It’s no secret that enterprises today are having a tough time controlling data growth. What is less well known, according to Marshall Amaldas, senior research analyst at IDC, is how valuable and prudent it is to have a strategic data retention and expiration plan in place. Without one, four basic problems are likely to occur:

- Data is expensive to store and becomes more so over time. “Having older, unwanted data that’s not being accessed sitting on primary SAN storage is a big waste,” says Amaldas.
- Data is expensive and time-consuming to back up and recover. As it grows, the risk of downtime increases.
- Excess data hinders performance. “You may just have to throw more storage at it, which ultimately slows down overall performance,” adds Amaldas.
- Growing amounts of data increase potential risk and cost in eDiscovery and compliance.

The core of the problem is that unstructured data such as email, documents, spreadsheets, and multimedia are growing quickly worldwide. IDC projects a worldwide compound annual growth rate (CAGR) in disk consumption for unstructured data of 59 percent between 2010 and 2015.1

With these challenges in mind, more and more firms are being proactive about putting retention and deletion policies in place, Amaldas says. There are strong business and legal drivers that make the creation and implementation of data retention and expiration policies a top priority for IT leaders.

To put it simply, if a company has a workable and active document retention policy and deletes documents in compliance, it can convince a court during eDiscovery that the documents were not willfully destroyed. The well-known Zubulake v. UBS Warburg decision notes that even companies in continuous litigation are not required to keep every “shred of paper, every email, or electronic document and every backup tape—such a requirement would cripple large corporations.”2

As a result, if industry regulations or litigation hold requirements don’t specify otherwise, an enterprise need only keep electronic information as long as necessary for business purposes—and not any longer. Information not managed by a data retention policy and kept longer than necessary is not only costly, it’s risky.

In short, to delete is sweet, if done consistently and in compliance with regulations and a standardized retention policy. To make developing and enforcing that policy easier, CIO Digest gathered tips and strategies from IT decision makers and analysts. They make a compelling case to do so; the benefits include enhanced compliance, streamlined eDiscovery, and lower costs.
1. Automate a survey of your data
What data do you have? Where is it stored? Which applications generate it? Who owns it? How often is it accessed?

“These are important questions to answer when developing a document retention policy, and organizations can certainly do a better job of answering them,” says Amandas. He notes that typical IT teams devote resources to answering only a few basic questions about their data, such as which applications and workloads are growing the fastest. And they do so simply to monitor storage consumption with the objective of knowing when to add more disk so that they don’t unexpectedly run out.

Much more should be asked and can be learned, Amandas points out. There are technologies such as Symantec Data Insight and Symantec Data Loss Prevention that enable teams to automatically map the information in their enterprise, identify file types, connect users to departments and lines of business, and locate inactive, orphan, or sensitive data. “But it’s uncommon to find an organization having such deep insights,” he says.

2. Identify relevant business and compliance requirements
Besides mapping data, you need to identify the corresponding business and legal requirements for retention.

“This is a particular challenge for the City of Miami Beach, which has 33 distinct business units, including fire, police, human resources, finances, public works, and sanitation departments. To complicate matters further, each unit has different business and regulatory requirements.

In addition, the entire City of Miami Beach government must comply with Florida’s Sunshine Law, which specifies that members of the public who request access to documents need to be able to see them, unless there’s a security issue.

“The different retention requirements really become a jungle that you have to maneuver through,” says Nelson Martinez, IT division director, City of Miami Beach. “If you want to do the smart thing, you bring in professional services that have expertise in assisting organizations in records management. The guidelines haven’t really changed, and there is much that consultants have learned from repetition in meeting them.”

To streamline the ability to respond to the Sunshine Law and easily discover information, the City of Miami Beach uses an archiving solution, Symantec Enterprise Vault with its Discovery Accelerator feature, to archive both email and files. It also uses Symantec NetBackup for data protection. Since deploying the two integrated solutions, Martinez has been able to redirect three full-time-equivalent staff to more valuable projects. “I can’t afford to have resources taken out of the loop for a week or two just looking for email,” he says. “In a busy week, we can get up to 15 to 20 requests.”

Symantec Enterprise Vault deduplicates and compresses the email and files it archives. Symantec NetBackup deduplicates the backup copy of all data. The two solutions reduce storage consumption for the city. And because NetBackup integrates with Enterprise Vault, the IT team is able to define automatic, policy-based migration strategies.
Your Data Is Doomed Anyway

The most enduring human data on earth is rapidly leaving the solar system as you read this. Launched 35 years ago, the Mars Voyager I space probe flew past Pluto’s orbit in 1989 at 38,000 miles per hour. In 40,000 years, it’s expected to come within 1.6 light years of another star.

On board, as humanity’s ambassador, is a 12-inch gold-plated copper disk containing analog sounds and images that portray the diversity of life and culture on Earth. The record is nestled with a stylus and cartridge, ready to play. Contents include recorded greetings in 55 languages, pictures, and music selections such as Mozart, Beethoven, and Chuck Berry performing “Johnny B. Goode.”

Back on earth, no retention policy can make data truly last. The U.S. National Archives cites a life expectancy for optical media of between 5 and 10 years and for magnetic tape of about 1 to 25 years.

“If you want to keep information forever, it has a price tag on it,” says Ariel Pisetzkzy, CIO at 888 Holdings PLC. “The storage in the Library of Alexandria was burnt down, and we’ll never know what was on those scrolls. The same might happen to any organization.”

Pisetzkzy visited a holocaust archive in Israel recently that had 50-year-old microfilm from the U.S. government. “Some of it had already withered into dust, literally,” he says. “The manufacturer had promised it should last 100 years, but after 50 it’s nonexistent. The manufacturer is gone, as well. There’s no one to sue.”

For the business value 888 Holdings is realizing from its data retention and expiration policies, download the Business Impact Study at go.symantec.com/888-bis.
He continues: “We can tell each business unit that its storage is X percent of the total storage and it’s utilizing Y servers. They understand their underlying IT costs. And they never want to be the most costly application on the radar or the most costly business unit if they’re not number one in revenue.”

6. Shrink the problem with deduplication and archiving

If you think your data growth challenge is big, consider the situation at Diplomat Distributors, Ltd., one of Israel’s leading distribution companies for fast-moving consumer goods, personal care products, food products, and cosmetics. Its data grows at 30 percent monthly.

Fast growth threatens availability and disaster recovery, as IDC’s Amaldas points out. Diplomat is especially sensitive to these risks, as its supply chain operations are in motion 24×7 and any interruption can be devastating.

The IT team at Diplomat turns to deduplication to help minimize data growth. They use NetBackup as the heartbeat of a two-tier storage area network (SAN) and network attached storage (NAS) environment, supporting a coordinated data retention, archiving, and data expiration process.

NetBackup replicates differential data on a daily basis from critical systems—including SAP—to a second data center located adjacent to the primary data store. “Deduplication has reduced the volume of data stored by as much as 90 percent,” says Dan Shisman, IT infrastructure & security manager at Diplomat. “That has improved performance. Backups that previously took up to 72 hours are now being completed in only 14 hours.”

Once data movement is automated based on policies, there are other benefits besides retention management. Data becomes more useful, says Shisman. His team has configured NetBackup to move data productively between the production systems, a 1.5 terabyte development environment, and a 1.5 terabyte quality assurance (QA) system.

It works this way: data is initially backed up to the company’s EMC CLARiiON platform and then run off on a daily basis to an Oracle StorageTek tape library. “We back up and replicate the data,” says Shisman. “Then, we again use NetBackup to restore the environment from the production to the sandbox or the QA environment. This makes it very quick and easy to create a new test and development environment.” Process efficiency, retention management, and data protection go hand-in-hand.

NetBackup enables the team to keep certain data for a period of seven years, while other data has pre-built expiration policies, thereby allowing it to expire after several months to several years. This approach to data retention—in tandem with deduplication—is leading to a significant reduction in storage volumes for Diplomat. “We have not separated out the specific storage savings from data retention, but we can say that the combined blend of data retention policies and deduplication has saved approximately six terabytes of storage,” says Shisman.

7. Collaborate to review retention policies periodically

When a retention and expiration policy is set, says IDC’s Amaldas, “Organizations should have compliance, legal, and IT departments collaborate often to make sure the policy is refreshed. The policy should represent updated requirements from a particular department or new directives that may have been pushed from headquarters.”

Data that isn’t being used will be expired. We expect up to a 20 percent reduction in storage.

– Ariel Pisetzky, CIO, 888 Holdings PLC

Organization Profile
888 Holdings PLC
Founded: 1997
Location: Gibraltar, U.K.
Employees: 937
Website: www.888holdingsplc.com
Based on IDC survey results, Amaldas notes, “we are definitively seeing that more and more organizations collaborate closely on this topic—sometimes annually, quarterly, or monthly.”

At the City of Miami Beach, Martinez recognizes that requests for exceptions to the policy will occasionally come up. “There might be some pictures that a department wants to hold onto for historical purposes, for instance,” he says. “It’s expensive to discover without having the tools to go through and review the data.”

8. Get 100 percent payback in the first eDiscovery event

Why invest in implementing a document retention and expiration policy? The project can sometimes mean more than pay for itself in the first eDiscovery event, says IDC’s Amaldas. “Organizations sometimes choose to settle because the discovery cost is going to be prohibitively high,” he says. “It’s expensive to discover without having the tools to go through and review the data.”

9. Learn how to delete

Many organizations are overly cautious about deleting data, IDC’s Amaldas says. “Often when they’re moving from one legacy archiving system to another, they leave their old data in whatever system they were managing before,” he observes. “But they would apply the new policies to whatever new data is being generated.”

888 Holdings lets time destroy data that is due to expire. “We push it into a final backup, and then that backup is sent off,” says Pisetzky. “In a few years, those tapes deteriorate and wither into dust. And the organization has readily accepted this.”

At the City of Miami Beach, the IT team deletes data, but only after “a carbon-based life form signs off,” says Martinez. “You want to have someone review the data and create a disposition list.”

Solve by pressing here

Developing a retention and expiration policy can seem like a formidable project at the outset. “I can tell you that this problem is a 900-pound gorilla, and it doesn’t matter where you are in this world,” says Martinez. “Everyone’s dealing with the same issues and pains, regardless of country. Languages and regulations may differ, but we share the same problem.”

Luckily, however, it’s a problem that can disappear—once a document retention policy is in place—with the push of a button.

Did Duplication has reduced the volume of data stored by as much as 90 percent.”

– Dan Shisman, IT Infrastructure & Security Manager,
Diplomat Distributors, Ltd.