# Setting Retention Policy for Electronic Information: A Practical Guide

**Ferris Report #832. September 2010.**

Ferris Research, Inc.
One San Antonio Place
San Francisco, Calif.  94133, USA
Phone: +1 (415) 367-3436
www.ferris.com

# Table of Contents

# Executive Summary

A retention policy for electronic information defines how long the information should be stored until it’s deleted. Except for email, most electronic information is currently never deleted or is deleted in a haphazard manner.

Legal and regulatory pressure is gradually driving organizations to define and implement proper retention policies and procedures for electronic information. Large, highly regulated organizations are in the lead. But even in such organizations, the situation is a mess, with arbitrary retention periods applied to many types of information.

Determining and implementing retention policy for electronic material is difficult. It’s hard to determine the right policies because many groups are involved and the concepts are subtle. And it’s hard to implement the policies, for example because the computer support tools are severely inadequate.

This white paper presents practical recommendations for organizations wishing to define and implement retention policy for electronic information. It also describes the current state of retention policies so an organization can compare its status with that of its peers. The paper also presents our predictions for the retention situation in 2020.

The paper is unique in that it:

- Presents, for the first time, the overall picture of corporate electronic retention. Other work on electronic retention limits itself, narrowly, to email
- Provides practical and actionable advice on how to develop and implement electronic retention policy
- Provides planning insight, by formulating a clear vision of how electronic retention will take place in 2020

### *Credits to Interviewees*

We conducted a series of interviews to validate the recommendations and conclusions of this white paper. We especially want to thank the following people for their valuable insights (and their enjoyable conversation):

- Svend Frandsen, ComArchive
- Ken Courtney, Firestone Diversified Products
- Sue Trombley and her colleagues, Iron Mountain
- MaryKay Roberto, Mimecast
- Tim Haugen, Nationwide
- Ralph Ehlers, Roche
- Kon Leong and his colleagues, ZL Technologies

IRON MOUNTAIN®

### *About our Sponsor, Iron Mountain*

Iron Mountain Incorporated (NYSE: IRM) provides information management services that help organizations lower the costs, risks and inefficiencies of managing their physical and digital data. The company's solutions enable customers to protect and better use their information—regardless of its format, location or lifecycle stage—so they can optimize their business and ensure proper recovery, compliance and discovery. Founded in 1951, Iron Mountain manages billions of information assets, including business records, electronic files, medical data, emails and more for organizations around the world.

Iron Mountain's comprehensive Compliant Records Management Program (www.ironmountain.com/compliance/compliance-records-management.html) can provide the expert guidance and support you need to ensure compliance with evolving government and industry regulations while reducing costs. Even better, you'll protect your company by increasing control over critical records, including email. Visit www.ironmountain.com or follow the company on Twitter @IronMountainInc for more information.

*Text provided by Iron Mountain*

We're grateful to our sponsors for their financial support, without which this research would not have been possible. Their views on the topics discussed in this report were also of great value.

You may copy or freely reproduce this document provided you disclose authorship and sponsorship and include this notice. Ferris Research independently conducted all research for this document and retained full editorial control.

# Background

### *What a Retention Policy Is*

A retention policy is a rule that states how long electronic information should be kept before it's destroyed. Electronic information can take many forms, such as emails, text messages, word processing documents, spreadsheets, PDFs, and relational databases.

Common rules for electronic information include:

- Don't delete any electronic information, apart from obvious junk like spam. Keep everything forever.
- Delete everything after 60 days.
- Keep taxation material for seven years.
- Keep personnel records for two or three years after the employee leaves.
- Keep bridge design materials for 50 years.
- Keep contracts for four years.

An organization's retention policy typically consists of a series of such rules. Together they are known as a retention policy or, sometimes, retention policies. Our focus in this paper is the deletion of material at the end of its lifetime, but of course this does not detract from the importance of archiving such material in the first place.

### *Reasons to Define a Retention Policy*

The main reasons for defining a retention policy are:

- To reduce the costs of e-discovery in the event of a lawsuit or other legal action. Producing information in response to an e-discovery request can be extremely time-consuming and costly. If the material has been deleted, the costs of production are obviously reduced. Following a clear, communicated retention policy adds defensibility to the deletion of electronic information.
- To reduce the dangers of e-discovery. Minimizing the amount of electronic material an organization keeps means it has less material to produce during e-discovery—and consequently it is less likely to hand over incriminating evidence.
- To reduce the costs of managing the storage needed for email. For example, many email systems need to limit user mailboxes to 250MB or so. Offloading email from the mailbox takes time and technical support, and the hidden costs soon mount.
- To satisfy laws, industry regulations, or individual corporate policies requiring that information be deleted in certain conditions. For example, laws might require private information about former employees or customers to be deleted after a certain period of time.

- To facilitate mergers and acquisitions. If part of a business is being sold off, the purchasing organization needs to have a clear understanding of the information it is taking with it. Deleting information supports that requirement.

### *Records and Records Managers*

A corporate record can be defined as a document that chronicles some important business transaction. Examples are invoices, contracts, stock transactions, and employment agreements.

In that sense, a large proportion of corporate documents aren't records but rather things like draft proposals or emails discussing where to have lunch.

In our view, the concept of a record is somewhat fuzzy, and reasonable and informed people within an organization often disagree about whether a specific piece of information constitutes a record.

Until the turn of the century (the 21st century, that is), records were mostly paper and had to be stored for long periods. The repositories varied from office cupboards stuffed with storage boxes to underground vaults in former bunkers and mines.

The employees responsible for managing all this paper were known as records managers. Clearly, they liked retention policies because it meant they could periodically weed out material and reduce the volume of physical storage. When you're storing paper records, it quickly becomes apparent that you can't keep all that paper forever.

Since about 2000, far more information has been stored electronically than on paper. Many records now exist solely or partially in electronic form, and records managers see their work as administering IT systems rather than shuffling around aging cartloads of paper. Unfortunately, while paper-based records often have orderly disposal procedures, the destruction of aging electronic information is in its infancy.

### *Limited Automation*

Today, email is the main type of electronic information subject to retention policy. That's natural, because email is by far the most important type of material used in e-discovery. Email is a particularly insightful electronic medium as it documents the date, time, participants and content of a particular conversation.

However, most retention policies are not specific to email. They refer to the type of information rather than the electronic format of that information.

The availability of technology to automate retention policy is limited. Email is the most advanced general-purpose tool in this regard. Instant messaging systems may also have automatic archiving and retention logic, as may specialized line-of-business applications.

Beyond email and instant messaging, retention tools may be found for SharePoint spaces, the contents of enterprise content management (ECM) systems, and files in a file share. Social networks are beginning to be supported. Nevertheless, except for email, the implementation of a retention policy is commonly not automated. Deletion is simply left to users.

### *Classifications*

Classifications are short descriptions given to different types of information so that retention policy can be applied. For example, electronically stored information might be classified as “tax record,” “HR record,” or “contract.” Classifications can also indicate whether material needs to be kept for e-discovery purposes; for example, “on e-discovery hold.”

Classifications are important because they facilitate the automatic disposal of electronic material. For example, a computer can easily search for documents that are classified as contract, are over 10 years old and haven’t been accessed for five years, and then delete them.

Material can be classified by users, typically by selecting from a drop-down list. However, most classification has to be done automatically.

# Current Landscape

Few organizations have thought through their retention policies and then implemented them.

Organizations with more than 10,000 employees are the most likely to have proper retention policies. Among these, those in highly regulated industries—typically financial services, insurance, pharmaceuticals, and healthcare—lead the pack.

Medium-size organizations (those with 100 to 1,000 employees) generally have few retention policies. Legal firms are often an exception, with proper retention policies in place.

Small organizations (those with fewer than 100 employees) almost never have retention policies. The exception in both these classes are businesses in highly regulated industries, as mentioned above. Otherwise, the retention policies for medium-size and small organizations typically involve deleting email to keep mailboxes small and to make the system function efficiently.

### *Traditional: Delete Email After 30/60/90 Days*

Until about 2005, the retention policy for electronic information was generally along these lines:

- Delete all emails after 30 (or 60 or 90) days, unless the user has made a specific effort to keep them.
- Keep all other electronic material as long as users want, or as long as the computer application allows.

This policy was mainly adopted to meet email systems' need to keep mailboxes small.

### *Transitional: Delete Everything After 30/60/90 Days*

A few years ago, in response to the growing threat of e-discovery, corporate legal departments began to want as much material discarded as possible, for two reasons:

- Primarily, to reduce the costs of reviewing the e-discovery material requested by opposing parties
- Secondly, to reduce the risk that their organization will be caught with incriminating evidence during e-discovery

Thus, the following type of retention policy became common:

- Delete all emails after 30 (or 60 or 90) days, unless the user has made a specific effort to keep them.
- Delete all users files (such as word processing or spreadsheet files) after 30 (or 60 or 90) days, unless the user has made a specific effort to keep them or there is a specific reason to keep them.
- Allow line-of-business, customized applications to maintain their own retention policies; though, wherever possible, information should be discarded after 30 (or 60 or 90) days.

In short, the default in this scenario is to keep things for 30 or 60 or 90 days and then discard them. With the exception of email, however, few general-purpose tools are available to automate this process, and much of the deletion is left informally to users or computer administrators.

This policy approach is gradually falling out of favor because courts have become concerned about the deliberate destruction of material that might reasonably be needed in later e-discovery. Several cases, including Adams v. Dell, have targeted the foundation of this approach by suggesting that retention policies need to be accountable to all corporate stakeholders, including external parties.

### *Today: Keep Everything*

As blanket deletion policies have become suspect—and, in some countries, against the law—the new approach to retention is:

- Keep everything and never delete it, except for spam and other obviously useless stuff.

For many organizations, this approach has a big advantage. Setting retention policy is hard, as many readers of this white paper are acutely aware. Keeping everything until the retention policy can be defined and implemented seems a safe and relatively easy option, especially since the cost of storage is decreasing faster than the volume of material to be stored is growing.

### *Emerging: Delete After Disuse*

Some organizations are employing a new approach as a fallback position. If no retention policy has been defined for a type of material:

- Once the material has been kept for X years, then if it is not used for a continuous period of Y months, it should be deleted.
- Material on legal hold should however be retained until it's no longer on hold.

For example, all documents should be kept for at least three years. However, if it is not then used for any 12-month period, it should be deleted.

This approach is appealing. It’s simple and intuitive.

It is also reminiscent of what cemeteries in Paris do to prevent overcrowding. When a grave is 100 years old, the cemetery publishes notices stating that it plans to reuse the grave unless somebody objects. If nobody objects, it goes ahead and reuses it. If someone objects, it waits another 100 years and repeats the process. Eventually, and pretty quickly, most spaces are freed up. Out, out, brief candle!

Put another way, as anyone who has ever cleaned out a closet or an attic knows, it makes sense to ditch old things if they haven’t been used by anyone for a long time.

### *Summary*

Most electronic information, with the exception of email, is never deleted or is deleted in a haphazard manner.

Legal and regulatory pressure is driving organizations to determine and implement proper retention policies and procedures. Large, highly regulated organizations are in the lead. But even in such organizations, the situation is a mess, with arbitrary retention periods applied to most types of electronic information.

# How to Define Policy

This section presents our suggestions on how to define and implement retention policy.

## The Steering Committee

First, the key departments concerned with retention policy should form a steering committee.

### *Membership*

The steering committee should include representatives of departments specifically concerned about regulations compliance. Those departments vary from industry to industry. For example, manufacturing organizations typically need representatives from engineering and design. In banking, it's often retail and wealth management groups.

In addition to such line-of-business departments, the Legal/General Counsel, IT, Compliance, Risk Management, Records Management, Information Security, Information Privacy, and Finance departments should participate.

### *Fallback Lead: Legal*

If the lines of business are not available to participate, then Legal/General Counsel should be the driver. The Legal Department needs to define retention periods so the organization cannot be accused of deliberately deleting or altering material that could reasonably be expected to be requested during e-discovery.

### *IT Not the Driver*

It's a mistake to have the IT Department lead the retention policy-setting project. Rather, the user departments most affected by the laws and regulations should drive the process.

However, IT needs to participate, since it will provide systems that support the deletion of electronic material. It will also advise on cases where disposal tools are not available.

In general, IT cannot impose retention policy on departments. Departments must buy in to policy definition.

### *Records Manager/Administrator*

The committee should select one person whose task it is to ensure that the program is rolled out and complied with. That person should have good connections at the department level.

### *Extended Lifetime*

The steering committee needs to stay in existence over the long term:

- Policy implementation will be spread over years.
- As laws change and/or interpretations evolve, the organization’s retention policy will have to be revised.
- E-discovery regulations, in particular, are evolving quite quickly and require periodic changes to retention policy.

### *Top Management Support*

The steering committee’s task isn’t straightforward and requires resources from user departments. If top management does not support and help promote the steering committee, the project will run out of steam.

## General Approach

### *Network with Peers in Your Industry*

Retention regulations are often industry-specific. A good way to find out about them is to network with people from other organizations in the same industry. This is most commonly done at trade gatherings, although there may also be discussions on the Web.

Regulations are sometimes ambiguous, and it may not be clear what practical steps need to be adopted. So networking with peers is also useful to determine what constitutes reasonable, good-faith efforts to comply.

### *Define Policies at Functional, Not Departmental, Level*

When formulating policy, think about classes of information that are intrinsically important to the business. Avoid defining classes of information that are too tailored to a specific department.

So if a class of information spans several departments, seek a higher-level concept that spans the departments.

### *Define Policies Independent of Data Format*

When formulating policy, do it for types of information. Don’t worry about whether the information is embodied as emails or in spreadsheets: Think about information in all data formats.

### *Prioritize the Highest-Risk Areas*

Define the high-priority areas and plan to deal with them first. The main considerations are:

- Important information that may be prematurely destroyed
- Legal fines if material is not kept long enough
- Review costs during e-discovery if too much material is kept
- Fines from regulators and loss of shareholder confidence if material cannot be produced

Don’t expect to deal with everything immediately. It’s normal to define and implement retention policy gradually.

### *Let Countries and Departments Tailor Implementation*

Laws and regulations vary between counties and can be inconsistent. Therefore, global organizations should create an umbrella policy and then let individual countries apply it.

Countries and individual departments should also be given flexibility with regard to how policy is implemented. In some cases, for example, paper may have to be shredded; in others, material in a SharePoint directory may have to be deleted; and in yet others material in an ECM system may have to be deleted.

## Policy Formulation

### *Decide Classifications*

Draw up a list of document classifications that are needed to implement policy. For the most part, these will correspond to the types of information that need retention policies applied to them, such as tax records, human resources, contracts, final engineering drawings, intellectual property, or customer financial information.

Legal hold is another important classification. This indicates material that should not be deleted because it is subject to use in litigation.

### *Keep Policy Simple and Classifications Few*

The more policies and classifications an organization has, the harder it is to understand what's going on. This is especially so when human beings have to implement policy and make decisions about which classification to apply.

One interviewee commented that his large organization had been able to keep the retention rules to about 100. These were spread across many different departments in different countries so that a given department had to be concerned with only a handful of rules.

Similarly, avoid the temptation to define a large number of classifications. No matter the organization's size, it probably does not need more than 100 classifications. Try to keep things simple.

### *Cross-Industry Guidelines*

Here are some suggestions for specific retention policies that apply across many industries:

- Keep tax records for as long as the taxation body requires. In the United States, this is seven years.
- Because of the potential for litigation, keep personnel records for perhaps three years after the employee leaves. Note that privacy regulations may require destruction after some period of time.
- Keep real estate records, such as those relating to office rentals, for four years.
- Keep major contracts for 15 years, and other contracts for four years.

### *Consider Statutes of Limitations*

Statutes of limitations impose time limits to discourage unreasonable delays in lodging lawsuits or criminal prosecutions. These time limits vary from topic to topic, country to country, and state to state.

For example, in Ohio, the time limits for starting civil cases are 21 years to recover real estate; 15 years to sue on written contracts; six years to sue on oral contracts; two years for actions related to personal injury or property damage; and one year for libel, slander, malicious prosecution, false imprisonment, and professional malpractice. Most other types of lawsuits are subject to a four-year statute of limitations.

Statutes of limitations provide useful direction on retention periods. In sensitive areas, such as contracts and personnel files, organizations should keep material as long as is it could be needed in a lawsuit or prosecution.

### *Consider Delete-If-Not-Used-For Period*

For material without a defined retention policy, consider this fallback policy: *Delete material that is X years old and has not been accessed for a period of Y years*.

This approach makes sense to us. For example, getting rid of material that is five years old and has not been looked at for at least two years seems safe.

Organizational concerns about the quantity of material that may need to be reviewed for e-discovery argue for destroying material as quickly as possible. We think it is reasonable to get rid of material that is three years old and has not been looked at for at least 12 months.

## Procedures and Ongoing Maintenance

### *Decide What to Automate*

Once policy has been determined for the different types of information, decide which types of information are the most important. Then try to apply automated retention policies to them.

As mentioned, it’s often only practical to provide automated retention policies to email and perhaps instant messages. Vendors are slowly increasing the range of applications that can have retention policies automatically assigned to them. After email, the order in which an organization will automate retention policies is typically:

- Instant messages (if the organization is heavily regulated)
- SharePoint or other teamspaces
- Fileshares
- Social networks

### *Decide What Users Will Manually Delete and Classify*

Think about what type of material will require manual classification or deletion by users, and whether it's practical to have them do it.

Even with rigorous and ongoing training programs, manual interventions are likely to be performed poorly. Users forget to classify, or they apply the wrong classification. And they forget to delete material, or they do the deletions poorly (for instance, by only partially deleting information).

Thus, wherever possible, retention policies should minimize manual user intervention.

### *Define and Implement User Education*

Once policies have been defined, users must be educated on the policies and the reasons behind them:

- Much electronic material will only be destroyed if users take action.
- For compliance purposes, an organization must be able to show that it has made reasonable and competent efforts to ensure that retention policies are implemented.
- In the case of email, which is the main application in which disposal can be automated, it's important to educate users because otherwise they'll resist. PST files (local copies of a mailbox, maintained by the user) must often be deleted, and without proper explanation many users will be extremely unhappy. They get very attached to their email.

It may be good to organize the training similar to other compliance-oriented training such as training for occupational health and safety, or nondiscrimination in employment.

Computer-based training is often effective. Ideally, people should sign off that they have taken the course and understood it; annual retests should also be given.

### *Define and Implement Audits*

A regular audit process should be put in place:

- To ensure that material is being deleted according to policy and that users are familiar with the policies that apply to them
- To be able to demonstrate, for compliance purposes, that the organization has made reasonable and competent efforts to ensure that retention policies are implemented

### *Document Policies and Implementation Plan*

Rollout plans should be documented. If the organization is audited by internal or external teams, this show it is making good-faith efforts to develop and implement an effective retention policy.

# Retention in 2020

Confusion about setting and maintaining retention policy will likely continue for 10 more years. So as the dust settles, say in 2020, what will most organizations' retention policies look like?

Our best guess is that it'll be along these lines.

### *Technology Developments*

On average, the decrease in storage costs will continue to outpace the increase in volumes of electronic material. Note that this assumption could quite easily be proven incorrect. To illustrate, consider Moore's Law, which states that the number of transistors that can be put on a chip doubles every 18 months. However, it is becoming harder and harder to squeeze in more transistors, and Moore's Law is beginning to fail.

Organizations will be able to define a set of classifications and then apply a given classification in a standard way to many types of data structures: emails, database entries, voicemail messages, postings in shared workspaces, files, and so on.

It will be possible to define retention policy centrally, in a way that applies to many different types of computer applications and data structure. For example, "Keep tax information for seven years and then delete" will be able to be automatically applied to an organization's email, file, voice message, and its ERP systems.

A major driver today for retention policy is that legal departments need to have reasonably small search results that they can then whittle down by human review. Search technology will continue to improve, making it easier for legal departments to keep search sets to a tolerable size. Likewise, review systems will become more productive. The pain and cost of e-discovery reviews will become far less shrill than they are today.

Automatic classification technology—closely related to that of search—will also improve in accuracy and usefulness.

### *Policy Developments*

Users will become accustomed to manually classifying a small number of document types. However, most documents will not be manually classified because the process is error-prone and distracting to users.

Setting and maintaining retention policy is difficult and takes resources. Thus, most medium-size organizations (those with 100 to 1,000 employees) outside heavily regulated industries will tend to:

- Apply retention policy to only limited types of information, such as tax records, human resources, contracts, final engineering drawings, intellectual property, or customer financial information. Often, the policies will comply with regulations requiring the deletion of private information.

- Adopt a default disposal process for material that does not have a retention policy—something along the lines of *Delete material that is over three years old if it has not been accessed for at least 12 months*.
- In short, delete material unless there is a specific, conscious effort saying it should be preserved.

Smaller organizations (those with fewer than 100 employees) will often weed out old material periodically by using search tools to define sets of information that can potentially be discarded and then applying rules such as *Delete material that is over three years old if it has not been accessed for at least 12 months*.

Large organizations will have complex retention policies. They have the IT and compliance resources to provide the necessary support and enforcement.

### ***Big Brother Will Be Watching You***

Notwithstanding good-faith efforts made by most organizations to delete personal information, a lot of it will continue to exist. It's just too hard to control this information. It's widely scattered and in unpredictable places. For example, individuals often keep information about their former work environments on their personal computers and handheld devices.

People are leaving ever-richer traces of their activities, and soon it will be possible for Big Brother to identify your hobbies and travel movements and much else about you through Internet-based forensics. Government organizations will also probably have access to private corporate data, whether surreptitiously or by more open means.

Something should be done about this. Technology can help to some extent by using encryption to protect privacy. But the larger solution is probably through laws requiring that private information only be used in certain ways.

*Author: David Ferris*
*Editor: Mona Cohen*

# About Ferris Research

Messaging. Collaboration. Compliance. Ferris Research analysts bring more experience in these areas than any other firm. Period.

Major areas of interest are email, archiving, e-discovery, information leak prevention, unified communications, instant messaging, SharePoint, and mobile communications. We help:

- IT staff evaluate, implement, and maintain these technologies
- Vendors understand the marketplace and its technologies; explain their products or services to the marketplace; and find strategic partners, raise funds, or sell their company
- Investors find and evaluate investment opportunities

We've been in business since 1990—longer than any other analyst firm in our field:

- Clients include many of the world's largest organizations as well as computer vendors from major corporations to small startups.
- We have published more than 200 formal reports and 1,100 short bulletins.
- Our news service has approximately 10,000 readers and covers more than 2,000 highly specialized announcements annually.
- Our research team shares many decades of experience in our core competencies.

In short, our technology and industry depth helps you understand today's products, where they've come from, where they're going, and their value.

Ferris Research is located at One San Antonio Place, San Francisco, Calif. 94133, USA. For more information, visit www.ferris.com or call +1 (415) 367-3436.

### *Need Help Developing Your Retention Policy?*

For information on how we can help your organization build and maintain a retention policy for electronic material, contact David Ferris on +1 415 367 3436, or david.ferris@ferris.com.

### *Free News Service*

Ferris Research publishes a free daily news service to help you keep current on archiving, compliance, e-discovery, and related topics. The newsletter also covers messaging and collaboration. To register, go to www.ferris.com/forms/newsletter_signup.php.