RE-ENVISIONING THE RETENTION SCHEDULE

How to Build a Software-Ready Retention Schedule

Bruce Miller, IGA, MBA
RIMtech Consulting
ABSTRACT

Modern electronic recordkeeping software provides new capabilities and techniques for managing digital records that were not previously possible with physical records. These capabilities include such things as multiple retention rules per category, retention assignment based on document value, multiple retention triggers and types, retention override, automatic declaration of records, and more. Many of these new capabilities are possible due to the presence of metadata fields assigned to digital records. To fully utilize these new capabilities, the retention schedule must leverage known available document metadata, and call for new document metadata in support of retention. The schedule must be fully aware of the new capabilities and utilize them where appropriate. A retention schedule that fully leverages these new capabilities is referred to as a software-ready retention schedule. It is structured differently from a traditional retention schedule, allows multiple retention rules per category, leverages document metadata, uses multiple types of retention triggers unique to digital records, and explicitly specifies how case-based records are handled, as well as a number of other functional capabilities.

THE NEED FOR A NEW RETENTION SCHEDULE

Many organizations have decided to deploy a modern EDRMS (Electronic Document and Records Management System). The records administrators of such projects will soon learn that the retention schedule is the key building block underpinning a successful EDRMS.

EDRMS is a blend of two technologies. The first is a modern ECM (Enterprise Content Management) platform (which used to be known as document management). This platform forms a digital repository for all electronic records, and provides for advanced searching by content and metadata, security control, version management, workflow automation, and collaboration such as multi-author document editing, and much more. The second technology is recordkeeping capability, often delivered as a set of features within the ECM itself or as a third-party product added to the content management platform.

In reality, the retention schedule underpins both technologies. The retention schedule does more than just feed retention rules to the ECM platform—it actually greatly influences the configuration of the ECM itself. This is necessary for the recordkeeping component to do its job properly.

All modern EDRMS systems incorporate RBR (Rules-Based Recordkeeping) to some extent. RBR is an approach to electronic recordkeeping that automates the recordkeeping functions the end user would normally have to carry out. These functions include identifying which documents are records, when to declare documents as records, and how to classify the documents against the retention schedule. A full and proper EDRMS deployment that
fully utilizes RBR capability automates all these end user recordkeeping functions. End users have absolutely no role to play in the declaration or classification of any records. They simply operate the system as an ordinary everyday ECM, without thinking about records management whatsoever. Thanks to RBR however, in the background documents are being declared as records and are being properly classified against the retention schedule, even if the user is blissfully unaware of this.

Modern electronic recordkeeping software can carry out retention and disposition in ways most records professionals may not have even heard of. Because the records are digital, administrators have more document-level information to deal with and can leverage that information to do more granular, more sophisticated, and more flexible retention and disposition. For example, they can apply retention based on the value of documents, they can apply multiple retention rules to a single category, even different types of retention rules within the same category. The software has these amazing retention and disposition capabilities; however, the records administrator must tell it what they want it to do. And that's the job of the retention schedule. If we know what the recordkeeping software is capable of in terms of retention and disposition, then we can write a retention schedule to take full advantage of these powerful new capabilities. A retention schedule that leverages these retention and disposition capabilities is referred to as a “software ready” retention schedule.

Traditional retention schedules were written without any knowledge of the capabilities of modern recordkeeping software. If you use a traditional schedule within a modern EDRMS, the software won’t be able to utilize any of the advanced retention and disposition capabilities it delivers. Furthermore, it will severely curtail the ability to fully utilize modern RBR automation techniques. A software ready retention schedule however is written with the assumption that it will be used within an EDRMS, and will take full advantage of the advanced retention and disposition capabilities of the software. Any well-written software ready retention schedule can be used with any modern recordkeeping software, regardless of brand.

Figure 1 shows an excerpt from an admittedly dated retention schedule. It shows the title, description, and a very simplistic retention rule for each category. But it is real, and it is in use today.
Figure 1 - Traditional Retention Schedule

Figure 2 below shows just how different a modern software ready retention schedule looks in comparison to a traditional schedule. A traditional schedule typically is just a long list of activities and retention rules and citations. A modern retention schedule shown in figure 2 however has three different but interrelated components. More on this later.
In this paper, we will explain how the retention schedule plays a pivotal role in the overall configuration of a modern EDRMS, and highlight the characteristics of a software ready retention schedule.
THE ROLE OF THE RETENTION SCHEDULE

Figure 3 shows what a modern EDRMS looks like conceptually. There are three “layers” to an EDRMS:

**The retention schedule**
The software ready retention schedule. This will be divided into case categories and administrative categories. On the left side are two administrative categories (operator rounds, and employee onboarding). On the right are two case categories (union grievances, and safety audits).

**ECM structure**
Often referred to as “information architecture”, the ECM structure consists of all the so-called “libraries”, or places that documents can be stored. Different ECM products have different names for storage locations. Storage locations can be called libraries, folders, cabinets, etc. ECM structure also consists of the metadata, fields of information permanently stored with each document placed in each storage location. There is more to ECM structure than just libraries and metadata, including such things as versioning, security and collaboration, etc. But for now, we’re only concerned with libraries and metadata.

**RBR rules**
RBR rules refer to the rules created within the recordkeeping software to automate the recordkeeping processes, namely declaration (which documents are declared as records and when), and which retention rules in the retention schedule get applied to which locations in the ECM structure.

Done properly, the retention schedule massively impacts the ECM structure. Each category in the retention schedule translates to a library in the ECM structure. This library is where users will store documents for that particular category. Each category in the retention schedule forms one library in the ECM structure. Both the category and the library bear exactly the same name. Case categories require that the library be subdivided into “cases”,
or containers, one for each case. This allows us to group records of each case together, separate from and independently of all other cases.

At the top of the pyramid lies the recordkeeping software and its RBR rules. This is where you define declaration rule such as “if library = “operator rounds” and approved = “yes” then declare”. Retention rules also get defined here, such as “if library = “operator rounds” retention equals true document date +5 years”. The rules need to know what the library names are, and what metadata it can work with.

As you can see, the retention schedule forms the base upon which the ECM is structured. This in turn allows the RBR rules to execute against that structure, as shown in figure 3.

**CASE RECORDS**

The retention schedule must differentiate between case and so-called “administrative” record categories. Each category in the retention schedule therefore is either a case
category or an administrative category. In most organizations today, about 60% of all records belong to case categories. The best way to understand case records structure is with the help of an example. Suppose you have 1000 contracts in existence at any one time. Each contract has a contractor name, the contract value, an expiration date, a contract type, etc. This data will not change among all the documents in any given case. Each contract theoretically could have an expiration date different from those of all other contracts. All contracts would have a single retention rule similar to “keep five years after contract end date, and then destroy”. Although there is only one single rule applied to all 1000 contracts, that single rule has 1000 different trigger dates, i.e. 1000 different expiry dates. The recordkeeping software must therefore track each of these 1000 dates.

Let’s look at this from the perspective of an EDRMS end user. A user has a document related to a particular contract. The document may be an email suggesting several changes to the draft of the contract. The user must specify which of the 1000 contracts the document is related to. How is this accomplished? The user must have a way to choose from among the 1000 contracts. How this is done can vary among different ECM systems but the most common would be a simple drop-down list of all 1000 contracts, as shown in figure 4. Each contract has a unique name, and the user must select one of the 1000 contracts. The ECM will have a library known as “contracts”. That library will be further subdivided into 1000 case containers, each bearing a unique name of one of the 1000 contracts. This is a good example of how the retention schedule shapes the ECM structure. The two have to work in concert, and only then can the RBR rules be applied to the records within these libraries.

![Figure 4 - Contract Selection](image-url)
RETENTION SCHEDULE STRUCTURE

A modern software ready retention schedule is recorded in a spreadsheet. There are 2 reasons for this:

1) **It is machine-readable.** All the elements of the retention schedule including all categories and RBR retention rules can be read by modern electronic recordkeeping software and imported directly into the ECM and or the electronic recordkeeping software itself.

2) **Better presentation.** In a spreadsheet we can group things by business unit, or by department. We can apply filters to various columns to examine subsets of the schedule. We can use automatic numbering to number categories. Compared to a written document, it is a better environment for developing the schedule, revising it, and presenting both to machines and to humans.

It does not matter which proprietary spreadsheet format you use (Microsoft Excel, Google Sheets, etc.). The examples we use in this report utilize Microsoft Excel. The retention schedule is a workbook consisting of multiple worksheets.

The retention schedule consists of 3 major components:

**Categories**

A worksheet containing all the categories for each business unit within the organization. Each category is named, numbered, and has a retention rule. Where there is more than one retention rule for the category, only one retention rule is shown and all of the category’s retention rules are listed in the MRR worksheet.

**Cases**

A worksheet containing the details such as naming nomenclature for each case across all business units.

**MRR**

A worksheet containing retention rules for each category that has more than one retention rule.

The first worksheet is a summary of primary business functions, as shown in figure 5 below.
Figure 5 - Primaries

In this worksheet, the code column heading is a short acronym for each of the business functions. The function is the name of the function. Number indicates a sequential number assigned to each of the primary business functions. The description is a detailed description of the function. Each row in this worksheet constitutes a different business unit grouping within the organization, often called a department or section. Each row in this worksheet has a corresponding worksheet of the same name.

CATEGORIES

Figure 6 shows a worksheet for one of the business functions, in this case the clerks department.
Figure 6 - Business Unit Categories

Each spreadsheet row is a single category. White rows are administrative categories, usually with simple time-based retention rules, and green rows indicate case categories that are subdivided into cases. This report does not allow for a comprehensive treatment of all column headings, therefore we will highlight only the key headings from figure 6. The primary headings are as follows:

**Secondary**  
Short title of the category.

**No.**  
Sequential unique number of the category.

**Description**  
Detailed description of the category.

**MRR number**  
Indicates that there are multiple retention rules for this category. The rules appear in the MRR worksheet. Each batch of rules unique to that category are uniquely numbered.
BR  Business retention. Retention needed by the business, not the retention period specified by legislation.

**Trigger**  Either the document metadata field, or the case metadata field used to trigger the retention period

**Type**  One of the 5 retention types (explained later in this report).

**Unit**  Unit of time measure, typically years.

**Disp. Action**  Disposition action. What will happen to the records at the end of their lifecycle. Typically delete, keep permanently, review, transfer.

**CASES**

Figure 7 below shows the worksheet used to define the details (structure) of all cases.

<table>
<thead>
<tr>
<th>Category Name</th>
<th>Case Example</th>
<th>PRI</th>
<th>Case Naming Convention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Requests, HIPAA</td>
<td>Request 1: 01/01/2020</td>
<td>M</td>
<td>Request Name, M, 36, Request Date, M, 8</td>
</tr>
<tr>
<td>Accidents, Vehicles</td>
<td>Employee 1: 01/01/2020 Description 1</td>
<td>FLEET</td>
<td>Employee Name, M, 36, Accident Date, M, 8</td>
</tr>
<tr>
<td>Accounts Payable</td>
<td>2019, 2020</td>
<td>FIN</td>
<td>Fiscal Year, M, 4</td>
</tr>
<tr>
<td>Accounts Receivable</td>
<td>2019, 2020</td>
<td>LIB</td>
<td>Fiscal Year, M, 4</td>
</tr>
<tr>
<td>Acquisition, Facilities</td>
<td>Facility 1: 2020</td>
<td>FAC</td>
<td>Facility Name, M, 48</td>
</tr>
<tr>
<td>Administrative Guidelines and Procedures</td>
<td>Procedure 1 Rev 1</td>
<td>COMMON</td>
<td>Procedure Name, M, 48, Revision Number, M, 12</td>
</tr>
</tbody>
</table>

**Figure 7 - Case Structure**

The purpose of this worksheet is to specify the naming convention for each case within each category that is designated as a case category. Each case within a category must be uniquely named from all other cases in that same category. Some ECM systems have severe limitations on the length of a container name. A container is what the ECM uses to group together documents that are related. In some ECM systems this is called a folder, a cabinet, or a document set, etc. We will refer to these with the generic name “container”. We specify a case naming convention of 3 parts. Each part will specify a name, whether it is
mandatory or optional (M/O), and a maximum allowable number of characters for that part of the name. The column headings are as follows:

<table>
<thead>
<tr>
<th>Category name</th>
<th>The name (title) of the category.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case examples</td>
<td>Fictional examples of how the naming would appear for each case.</td>
</tr>
<tr>
<td>PRI</td>
<td>The primary business function underneath which the category falls.</td>
</tr>
<tr>
<td>Name</td>
<td>The name of that part. The system administrator names the container with a suitable name that matches the particular case, but this column shows what the name is supposed to consist of.</td>
</tr>
<tr>
<td>M/O</td>
<td>Either mandatory (M) or optional (O).</td>
</tr>
<tr>
<td>MAX</td>
<td>Maximum allowable number of characters.</td>
</tr>
</tbody>
</table>

**Multi-Rule Retentions**

This worksheet contains one row for each component of a retention rule for each category that specifies more than one retention rule. These rules can be read by machine directly into most modern recordkeeping software. This spreadsheet can also be easily manipulated so that the column headings and the order of the columns are in the particular order that the recordkeeping software requires. Column headings are as follows:

<table>
<thead>
<tr>
<th>MRR</th>
<th>A unique sequential number that identifies the batch of rules unique to a given category. Each row will have the same number for all rule components in a given category.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI</td>
<td>The primary business function that the category falls under.</td>
</tr>
<tr>
<td>Document field</td>
<td>The document metadata field that triggers the retention rule.</td>
</tr>
<tr>
<td>Value, document</td>
<td>The value of the document metadata field necessary to trigger the rule.</td>
</tr>
<tr>
<td>Case field</td>
<td>The case metadata field that triggers the retention rule.</td>
</tr>
<tr>
<td>Value, case</td>
<td>The value of the case metadata field necessary to trigger the rule.</td>
</tr>
<tr>
<td>Name</td>
<td>The external trigger name that triggers the retention rule. Typically from an external source such as a corporate database.</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Value</td>
<td>The value of the external trigger necessary to trigger the rule.</td>
</tr>
<tr>
<td>REL</td>
<td>Related. A Boolean logic operator that relates this rule component to the following rule’s component. Examples are AND, OR, NOT, etc.</td>
</tr>
<tr>
<td>Type</td>
<td>Retention rule type. Retention rule types are itemized later in this report.</td>
</tr>
<tr>
<td>Period</td>
<td>Retention period.</td>
</tr>
<tr>
<td>Unit</td>
<td>Unit of time measure, usually in years.</td>
</tr>
<tr>
<td>Disposition</td>
<td>Action carried out at the end of the lifecycle, e.g. transfer, permanent, etc.</td>
</tr>
</tbody>
</table>

**Retention Schedule Characteristics**

Here we will examine the five core structural characteristics of a software ready retention schedule. They are:

- **Multiple Retention Rules** The ability to have multiple retention rules, and multiple types of retention rules, for any given category in the schedule.

- **Value-Based Retention** The capability to base retention periods on the value of specified documents within the category.

- **Published Documents** A method of handling documents with an indeterminate retention period.

- **Retention Over-Ride (ROR)** The ability for an end user to override an assigned retention rule.

- **Continuous Over-Write** A means of dealing with records that are being continuously revised and updated.
**MULTIPLE RETENTION RULES**

Traditional retention schedules by and large allow only a single retention treatment for each category. That retention treatment, or rule, can be time-based as in “delete after 5 years”, or case-based as in “delete two years after end of investigation”. With the first rule each document qualifies for destruction is it reaches five years of age. Disposition is accomplished on a document by document basis. In the second rule all the records in any one case are qualified for disposition two years after the case has ended, i.e. the investigation is over. In both of these instances, there is one single retention rule that applies to all the records in that particular category.

Modern electronic recordkeeping software however allows us to support not only multiple retention rules for any one category, but also different types of rules within a category. Each retention rule type refers to a different approach used to calculate the eligibility for disposition. Inside the software, the retention type invokes a different algorithm that determines how the retention is calculated. Different software products offer a different selection of retention types. Some offer more retention types than others. And a given retention type in one product may be similar in function to that of another product, but will be named differently. The following table shows the five common retention types found across most software products:

<table>
<thead>
<tr>
<th>Type</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>Time Based (based on document's AGE)</td>
</tr>
<tr>
<td>D</td>
<td>Document-based (based on a document metadata field property)</td>
</tr>
<tr>
<td>E</td>
<td>Event Based (For case records, or external defined events)</td>
</tr>
<tr>
<td>R</td>
<td>Relationship-based (for Supersedence)</td>
</tr>
<tr>
<td>O</td>
<td>Over-Write. A document that is continually added to, over-writing prior changes, e.g. a tracking list or database. Must never be immutable, will never be deleted.</td>
</tr>
</tbody>
</table>
There are plenty of real-life situations that call for multiple retention rules within a given category. Below are some common examples:

1. Executed copies of agreements must be kept much longer than drafts and supporting or ancillary documents related to the contract.
2. Legislation specifies a different retention period applies if a document is referring to a person below a certain age.
3. In engineering projects, each type of document within the project has a different lifetime and value for retention purposes.
4. Approved documents are to be kept for longer than those that have not been approved.
5. Minutes and agenda of formal meetings are typically kept permanently, whereas the remaining records related to that meeting can be discarded.
6. The retention period of certain records can vary depending on the outcome of the business process. For instance, records related to the acquisition of a company specify that certain due diligence records are to be destroyed immediately if the acquisition fails to close, but if the acquisition is completed successfully, they are to be retained for X years.
7. Policy. Documents related to the policy can be discarded after a few years, whereas the official “published” policy that was put into effect remains indefinitely until superseded.

In any modern software ready retention schedule, it is common to have multiple retention rules applicable to as many as 80% of all categories in the schedule. Let’s take a look at a real-life example of a category in the retention schedule that requires multiple retention rules. Under human resources we have an activity (category) called “Credentials, Employee and Apprentice”. This is used to store all records related to the credentials needed by employees and apprentices, such as for operating vehicles with air brakes, handling hazardous materials, firefighting, or emergency medical services. There are three retention rules for such credentials, based on various applicable legislation as shown:

1. If **Hazardous Materials** = Yes, retention = Credential Expiration date + 50 years then destroy
2. If **Business Unit** = Fire or EMS, retention = Credential Expiration date + 8 years then destroy
3. If **Hazardous Materials** = No .and. **Business Unit** .not =. Fire or EMS, retention = 5 years

Let’s look at what these three rules really mean. The first rule states that if the credential is for hazardous materials, the records relating to that credential must be kept for 50 years then destroyed. Rule two states that if the record belongs to the business unit fire or the business unit EMS, and regardless of what type of credential it is, records relating to these credentials must be kept for eight years after the expiry of the credential, then destroyed.
Rule three looks rather complicated and technically it is somewhat complicated however its meaning is inherently simple. Rule 3 simply says that all other credentials are to be kept for 5 years then destroyed. This would apply to all credentials that are not hazardous materials, and are not within the business unit fire or the business unit EMS.

The recordkeeping software must have a way to know which rule to apply to which records within this category. It will rely on metadata to tell it what it needs to know. We need a document metadata field called “hazardous materials”. The default value will be NO. If, however the user enters YES in this field, that triggers rule 1 for that document. We need a second metadata field called “business unit”. If this field contains either “fire” or “EMS”, rule 2 will be applied to that document. Rule 3 will be applied to all remaining documents in that category.

This is an excellent example of how the retention schedule drives ECM structure. The retention schedule specifies the three different variations of retention treatments necessary for this category. It explicitly specifies the metadata fields needed in the ECM structure. As long as those metadata fields exist, and users use them, the retention rules will be applied correctly. Obviously these three fields must be mandatory, as the RBR retention rules depend on the values in these fields to do its work.

Figure 8 shows how these three retention rules are expressed in the retention schedule. The retention schedule is a spreadsheet consisting of several columns from left to right.

<table>
<thead>
<tr>
<th>Secondary</th>
<th>No.</th>
<th>C</th>
<th>Description</th>
<th>V</th>
<th>P</th>
<th>S</th>
<th>E</th>
<th>N</th>
<th>MRR No.</th>
<th>BR</th>
<th>Retention Rule</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credentials, Employee and Apprentices</td>
<td>170</td>
<td></td>
<td>The process of ensuring qualified employees and apprentices have a valid credential in effect for specified work that requires such credentials. Records include credentials, renewals, test results, tracking logs, and all correspondence related to the credentials and their expiration/renewal.</td>
<td>N</td>
<td>Y</td>
<td>C</td>
<td>100.1</td>
<td>5</td>
<td>True Document Date</td>
<td>T</td>
<td>Years</td>
<td>Delete</td>
</tr>
</tbody>
</table>

**Figure 8 - Three Retention rules in a single category**

The column "secondary" shows the title of the category. The column “MRR number” indicates that this category has multiple retention rules. MRR number 100.1 will show the details of the rules. Meanwhile, the column BR or business retention shows 5 (years). This is the default rule of five years then destroy, as needed for rule 3. However the column heading MRR shows rule number 100.1, which refers to the entire set of rules for this category. Let's take a look at the details of the retention rules for this category. Referred to figure 9 below.
There could be hundreds or even thousands of rules in this worksheet. In this category however there are exactly 8 rows which act together to form the three unique retention rules for this category, rows 79 through 86 inclusive. Each electronic recordkeeping software product has differing capabilities and limitations for multiple retention rules. Furthermore, each product has a slightly different approach and nomenclature to how the rules are expressed and documented. The example we see in figure 9 is a neutral expression of the three rules that should apply to most modern recordkeeping software products. They would likely have to be adjusted to suit any one particular software product.

In row 79 we define the first rule. The rule is triggered by the document field “hazardous materials”, and the value in the field must be “yes”. The retention rule type is T (time-based), the retention period is 50 years, and the disposition action is delete. Rule 2 is a little more complicated. In rows 80 and 81 we take care of the situation where the business unit is “fire”. In rows 82 and 83 we take care of the same situation, but where the business unit is “EMS”. In row 80, under the column heading REL (Related), we enter the Boolean operator AND. This simply means that the condition in row 80 and the condition in row 81 must both be met in order for the action to take place. In row 81, we specify that there must be a date in the metadata field “credential expiration date”. Hence if the business unit is “fire “and there is an expiration date, the document will be kept for eight years after the date specified in the field “credential expiration date”. Note the retention type is D, which tells the software to trigger the retention period on the date field known as “credential expiration date”. Rows 82 and 83 accomplish the same thing but for the business unit called “EMS”. Rows 80 to 83 together are necessary for our retention rule 2.

Rows 84 to 86 form our retention rule 3. Row 84 specifies that the field “business unit” must not contain the word “fire”. In row 85 we specify the field “business unit” must not contain the word “EMS”. In row 86 we specify that the field “hazardous materials” must contain the value...
“NO”. Once these 3 criteria have been satisfied the document will be kept for 5 years then deleted.

This has been a deliberately complicated example, but it shows how we can make very sophisticated and complex retention rules. Modern electronic recordkeeping software is more than capable of handling these complex rules; however we have to explicitly tell the software exactly what to do. This will require the use of metadata within the rules, and it is imperative that the retention schedule specify the metadata needed to execute the rules. This metadata must then be built into the ECM. Only when the metadata has been constructed can the rule possibly work. Over the lifetime of the ECM it is imperative that these metadata fields not be disturbed, renamed, removed, or altered in any way. If changes are made to this metadata at any time, the changes must be communicated to the RIM professional so the retention rule can be adjusted accordingly, otherwise the rule will simply stop working.

**Value-Based Retention**

With today’s recordkeeping software we can assign retention periods based on the value of records within a category. We can assign longer retention periods to documents of higher value, and shorter retention periods to documents of lesser value. To do this we again rely on document metadata within the ECM structure. We will need a metadata field to differentiate documents of high value from those of lower value. There are many ways to do this that can involve a single metadata field or multiple metadata fields depending on the particular activity. For now however we will use a very common technique found in a number of organizations. Let’s suppose we have an activity (category) for “capital projects”. These are large capital-intensive engineering projects such as building roads, bridges, or buildings. Each project is a case within the category. Each case will store all the records related to that particular project through to the end of its life (which would be the project end date). Needless to say, there could be thousands, even tens of thousands of documents for each project. We can define a metadata field that can tell us the nature of each document. The nature, or subject of the document in turn can tell us the inherent value of the document for the purposes of assigning a retention period. A good example would be a document metadata field called “Document Type, Capital Projects”. This would be a mandatory field in the ECM library, so that every single document must have a value in this field. There would be a dropdown list of document types similar to that shown below:
<table>
<thead>
<tr>
<th>Document Types, Capital Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document Type</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Project Management</td>
</tr>
<tr>
<td>Contractual/Legal</td>
</tr>
<tr>
<td>Planning and Logistics</td>
</tr>
<tr>
<td>Reports, Draft</td>
</tr>
<tr>
<td>Reports, Final</td>
</tr>
<tr>
<td>Charter/Authorization</td>
</tr>
<tr>
<td>Meeting Minutes/Agenda</td>
</tr>
<tr>
<td>Technical Specifications</td>
</tr>
<tr>
<td>Drawings, Draft</td>
</tr>
<tr>
<td>Drawings, As-Built</td>
</tr>
<tr>
<td>Regulatory-and Compliance</td>
</tr>
<tr>
<td>Permits &amp; Licenses</td>
</tr>
<tr>
<td>Contractor-related</td>
</tr>
<tr>
<td>Approval-related</td>
</tr>
<tr>
<td>Budget Related</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Retention triggers are as follows:

- **TDD**: True document date
- **EOL**: End of life (of asset)
- **EOP**: End of project

End users are forced to pick one of the 16 possible values for this mandatory field. Users would not normally see the trigger or the retention period when they select the document type. There is no reason why they couldn’t, but most users simply don’t have an interest in the retention periods. Below are some examples of how the retention rule was derived from the selection of the document type:

**Technical specifications** These records will be kept for 5 years after the end of life of the asset being constructed. If the asset is a bridge, technical specifications are necessary to keep on hand for the entire life of the bridge.
**Project management**  These records would include things such as project schedules, Gantt charts, and other documents related to the management of the project. The value diminishes quickly after they have been used, hence the retention period is the date of the document (true document date) +5 years, then destroy.

**Budget related**  Records related to the budget are to be kept for 5 years following the end of the project. These records are not needed to be held for the life of the asset under construction.

Not all document types necessarily need to have a different retention treatment from other document types. Note above that the two document types “contractor-related” and “reports, draft” each have the same retention treatment. In many modern EDRMS systems the document type is used to help end-users search and retrieve documents by their type. This is particularly useful where there are high volumes of documents, i.e. thousands or even tens of thousands of documents. The document type field makes it easier to find the document of interest. We can take advantage of that to assign appropriate retention periods to each document type.

Figure 10 shows how these retention rules will be entered into the retention schedule itself, in the MRR (multiple retention rule) worksheet.
**Figure 10 - Document Type Retention Rules**

Note there are two different retention types among the 16 retention rules. The rules in each of the spreadsheet rows 101, 104, 114, and 116 each use retention type T (time-based retention). The remaining rules use retention type E (event-based retention), except for the 3 rules in spreadsheet rows 105, 106 and 107 which call for permanent retention. Retention type E specifies that the trigger date is some event date. On spreadsheet row 108 the event date is the end of life (EOL) of the asset. In row 109 however the event date is the end of project (EOP).

This approach to value-based retention is generally useful when you have a very high quantity of records within a given activity (category). This approach offers 2 distinct benefits:
1) **Improved document retrievability.** Users can search for documents based on the type of document.

2) **Better retention granularity.** Documents of low continuing value are destroyed early, and documents of higher, more persistent value are kept longer.

Once again, it’s important to point out the criticality of metadata in the EDRMS. This technique would not be possible without properly defined metadata, in this case the “document type” field. Well-defined and carefully considered metadata is a key to the success of any ECM project, and is equally important for recordkeeping automation.

**PUBLISHED DOCUMENTS**

Certain types of documents have an effective retention period of “indefinite”. This usually means that the document is to be kept until it has been superseded by a newer version. The document is kept for an indefinite period of time until it is replaced with that newer version. Some examples may include:

<table>
<thead>
<tr>
<th>Policies</th>
<th>A policy, such as an email usage policy, is in effect until replaced by a newer version of the policy.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard Operating Procedures</strong></td>
<td>Standard operating procedures are often documented for things such as fire alarm drills, confined space entries, diagnostic test processes, plant operating and testing procedures, etc. Such procedures remain in effect, and must be followed, until they have been replaced with a newer version.</td>
</tr>
<tr>
<td><strong>Training materials</strong></td>
<td>Training materials have been developed for a particular training course. These materials are used to deliver the course as often as necessary. Eventually, these training materials will be replaced with a newer version. The retention period for the original training materials is <em>indefinite</em>, until superseded with a newer version.</td>
</tr>
</tbody>
</table>
Many plans are in effect until replaced with newer versions, such as annual operational plans, emergency plans, corporate strategies, etc. Sometimes plans are replaced on a scheduled cycle such as annual or every 5 years. However, in many cases a plan is in effect until it is replaced with a newer version, and it is impossible to predict when that newer version will arrive.

We refer to such documents as **published** documents. A published document is simply one that is “in effect” until superseded. The document is “in play” so to speak, or “in force”. We must not destroy these documents while they are in effect. Once they’ve been superseded, we can then apply retention. After the date they were superseded, we can then delete them. The term “published” is simply a convenient moniker, it is not necessary to use that particular word. Within any given category where a published document is being developed, there will be many more documents than just the published document itself. Suppose the published document being worked on is a policy. There will be many drafts of the policy. There will also be many emails with directives and instructions and comments related to the development of the policy. There will be many reference documents including financial documents, legal briefings, and any manner of supporting or ancillary documents. Of all of these records in that category, we only need to apply the supersedence process to the actual policy itself (which was put into effect). We can apply a different retention rule to the remaining documents. The remaining documents do not have an indefinite retention. They can be disposed of in a fixed period of time, or a certain period of time after the published document has been put into effect. Either way, we need a way to distinguish the published documents from those that are not. To do this we use a document metadata field called PUBLISHED (Y/N).

To handle supersedence in a modern EDRMS, we use a combination of the following four metadata fields:

- **Version**: The version of the document in question. Versions can be one of many forms, such as a sequential number, date, or even a season (summer, fall, etc.).

- **Superseded date**: The date that a document was superseded by a newer version.
Effective date  The date that a newer version of a superseded document was put into effect.

Published  A document with an indefinite retention period (until superseded). This distinguishes a document in the category that requires supersedence retention treatment from those that do not (drafts, commentary, supporting or ancillary documents).

The supersedence process is shown in figure 8.

Version 1 was published, or came into effect on January 10, 2018. On June 11, 2018 however version 2 was approved and took effect on that date. Hence the superseded date of version 1 became June 11, 2018, and the effective date of version 2 was also June 11, 2018. On December 10, 2018 version 2 was superseded by version 3, which became effective on December 10, 2018 version 3 has no superseded date as it has yet to be superseded by a newer version. Each of these 3 documents would have the value YES in
Suppose we have a category called “Policies, Corporate” with the following retention rules:

If Published = Yes, retention = Date Superseded + 5 years
If Published = No, retention = 2 years

Figure 12 shows how we would enter these two rules into the retention schedule in the MRR (Multi Retention Rules) worksheet.

In spreadsheet row 147 we indicate that the value YES must be in the field “published”. We enter “AND” in the REL (RELATED) column to show that there is a 2nd condition that must be met. In spreadsheet row 148 we specify that there must be a date in the field “superseded date”. The document will be destroyed 5 years after the date in the “superseded date” field. Note the retention type is D (trigger retention on a document metadata date field). In spreadsheet row 149 we deal with all the remaining documents, i.e. those that are not published. Here we simply keep these documents for 2 years, then destroy. The retention type T tells the software to destroy the documents 2 years following the true document date.

**Retention Over-Ride**

Every now and then in some categories, the business owner will request the ability to override the retention schedule and keep the document for a longer period of time. We refer to such an extension as a retention schedule “override”. The rationale for a business user to support this varies greatly but the following are some common examples of the reasons they may wish to override the retention schedule:

1) **Reference value.** A particular document may have an unusually long (persistent) value for future reference. It may be a rare legal precedent. It may be a technical specification or photograph of an extremely rare piece of equipment that is long out of date but still in service, and the document may need to be preserved as long as the equipment is still in service.
2) **Protective value.** A document may record something that could be used in the future to protect the organization from legal action, or serve to defend it in the event of any legal or regulatory challenge in the future. It may serve as evidence that the business owner feels should be kept well beyond the regular retention period, “Just in case”.

3) **Legal value.** Some legislation obligates an organization to keep relevant documents if there is a “reasonably foreseeable prospect” of legal action. You may suspect that this document or documents would be important in the event of future legal action against your organization.

4) **Historical value.** Records within a given category do not ordinarily contain anything of historical value. But for whatever reason, every now and then a document may be deemed to be historically significant, even though it was never expected to be. For example a photograph of the sod turning ceremony during the construction of a new facility may be included with the construction project documents, however the photo may be declared as historical, therefore you would wish to keep that particular photo permanently.

To allow an end user to override a retention period, you need a mechanism by which they can designate a document with higher retention value. This would be yet another metadata field. The field used for this override would commonly be called CRITICAL (Y/N), or something similar. The name of the field does not matter—it can be called anything you wish, as long as the user understands its purpose. We then define two separate retention rules for this category -- one where CRITICAL = NO, and one where CRITICAL = YES, as shown below:

- **If Critical = Yes, retention = 25 years**
- **If Critical = No, retention = 5 years**

Figure 13 shows how we would enter these rules into the retention schedule.

<table>
<thead>
<tr>
<th>MRR, PRI, RETENTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>162, Yes, 25 years</td>
</tr>
<tr>
<td>163, No, 5 years</td>
</tr>
</tbody>
</table>

**Figure 13 - Retention Rule Over-ride**

In spreadsheet row 162 we have a simple time-based retention rule where the value of the field “critical” = YES. Documents meeting that rule will be destroyed 25 years after the true document date. In spreadsheet row 163 we specify a retention rule of 5 years where “critical” = NO.

Below are some important considerations when implementing retention rule overrides:
1) Each category can have a different retention period for the override. For example, a category “financial audits” might have an override of 25 years, whereas a category “museum collections” might have an override of permanent.

2) This capability can be subject to abuse. Some users may have a proclivity to using it too much, on too many documents. The only defence against this is to educate your users, and to monitor the usage of the over-ride. We recommend that on a regular basis, perhaps monthly, you should run a report against the entire EDRMS to determine how often, in what categories, and which users applied the override. Monitor it frequently to ensure it’s being used in a healthy fashion and not being abused.

3) This override can be combined with other retention rules in any given category. The example below shows how the override can be applied to a category with supersedence rules:

   i. If Published = Yes, retention = Date Superseded + 5 years
   ii. If Published = No, retention = 2 years
   iii. If Critical = Yes, retention = 25 years

**Continuous Overwrite**

Unlike traditional physical documents, electronic documents can be modified on an ongoing basis over time. There are 3 distinctly different methods to modify a digital document:

**Save with a different filename**

Every time you modify the document, you save it with a different filename. This creates a separate document for each time you modified the document. Each is different from the other, and they bear a different filename from each other. Technically and legally, each change constitutes a different record. Each of these separate records can be declared and managed independently of the other.

**Save with the same filename**

You make a change to the document and save it without changing its name. This overwrites the previous version of the document with a new version that contains the changes. There is no record of the changes made to the document. There is no version that tells you how often it was changed or what the differences are between the versions. It is legally and technically one record, the content of which has
changed over time. This is what we refer to as a true “continuous overwrite” document. It is continuously being overwritten. The frequency with which it is modified is of no consequence, so don’t be put off by the “continuous” terminology. It is continuous insofar as all the changes are continuously being overwritten throughout the lifetime of the document.

Save and increment the version

In any modern ECM system, you can optionally turn on version management. Each time you save the document the system automatically increments the version number by one. The first time you save a document it will automatically have version 1. The next time you save it, it will have version 2, and so on. This allows you to go back in time and see every change made to the document. Legally and technically, each version is a record and can be managed as a record independently of all other versions. Some would say that the entire version series is a single record. Either way, in modern recordkeeping the versions and the changes to those versions should be preserved in accordance with the principle of records preservation.

Here we are concerned only with the 2nd of the 3 methods shown above – saving with the same filename. We refer to this as “continuous overwrite”. Some examples may include:

1) **Tracking logs.** A spreadsheet used to track attendance of students, phone calls, project changes, etc. The spreadsheet is updated periodically (daily, weekly, monthly) or on an as needed basis. Each time the spreadsheet is updated it is saved without changing its name.

2) **Databases.** You may have a database to track things such as assets, employee leave, or other information. Such databases might include Microsoft access, Oracle, or even a Microsoft Notepad document. The database is updated periodically and all the data is stored in a “database”. This database can consist of a single file or a set of related files, which we considered to be a record. The name of that database file(s) never changes, and is continuously overwritten as new data is added or modified.

3) **Notebooks.** A notebook can be an ordinary document such as a Microsoft Word document which is used to record operator notes, police notes, or anything else which is periodically and continuously updated. Microsoft has an innovative software application called OneNote which is specifically designed to record notes on a continuous basis in a single document. In essence, OneNote is a database of unstructured documents.
Such documents often constitute important records. But because they are being continuously modified and saved (i.e. continuous overwrite), we can neither make them immutable (locked them down and prevent deletion or modification) or delete them. So how do we deal with it in a retention schedule? Any given category can have one or more such documents. For example, a category such as “sales performance and tracking” might contain records of the sales quotas and targets of a sales team. Mixed within these records could be a tracking log – the spreadsheet that records and tracks aggregate data of the entire sales team over time. Because this tracking log is being continuously overwritten, we cannot make it immutable or delete it. In essence, we have to ignore it and leave it be. In this example we would define two retention rules as follows:

If **Continuous Overwrite** = Yes, retention = Ignore

If **Continuous Overwrite** = No, retention = 5 years

In the ECM, each document in this category must have a mandatory metadata field called “continuous overwrite”. The default value would be NO. For each tracking log stored in this ECM library (Category), the user must specify “continuous overwrite” = YES. Where “continuous overwrite” = YES, the software will ignore the document and not lock it down or apply any deletion to it. For all documents where “continuous overwrite” = NO, the document will be kept for 5 years then destroyed.

Figure 14 shows how we would enter that into the spreadsheet.

![Multi-Rule Retentions (MRRs)](image)

**Figure 14 - Continuous Overwrite**

In spreadsheet row 194 we specify that for all documents where document metadata field “continuous overwrite” = YES, we will use retention type O (ignore, no deletion). In spreadsheet row 195 we specify that for all documents where document metadata field “continuous overwrite” = NO, we use retention type T (time-based) and delete these documents 5 years after their true document date.
In some cases, tracking logs and similar databases are “rolled over” on a periodic basis. Suppose a tracking log is used to track sales performance in a given calendar year. A spreadsheet is updated continuously throughout the year. At the end of the year the spreadsheet for that year is left behind and a copy is made with a different name for the following year. This new spreadsheet for the next year is then continuously updated throughout the second year. This means that updating of each log ceases at the end of every calendar year. We can then apply retention to those logs. Suppose we had a category that contained tracking logs that were rolled over at the end of each year. If we applied normal default retention period of 5 years for example, that would be sufficient to capture the tracking logs. The tracking logs would be retained 5 years after the end of each calendar year, then destroyed. As a general rule of thumb, if the rollover period is less than the default retention period, the continuous overwrite treatment is not necessary.

**SUMMARY**

A fresh approach to the retention schedule is essential if you are about to deploy a modern EDRMS system. The schedule must be in a spreadsheet format, and contain detailed category descriptions, explicit cases and naming convention for case categories, as well as mathematically correct retention rules for all categories that require multiple retention rules. This allows you to leverage the full capabilities of the EDRMS software.

The retention schedule does more than simply specify retention rules – it forms the basic underlying structure of the ECM platform. Once the EDRMS is fully deployed, the ECM structure, the retention schedule, and the RBR automation rules all work together as an interconnected unit. Any change to either of these three components must be carefully coordinated so that the RBR rules do not break. This means the RIM professional must:

1) Substantially re-work the retention schedule
2) Heavily influence the ECM structure
3) Design and deploy automation records declaration and retention rules
4) Continuously monitor the overall system and ensure and changes are communicated and reflected in all three levels of the system
ABOUT THE AUTHOR

Bruce Miller, MBA, IGP is a world leading expert on electronic recordkeeping. He is an independent consultant, an author, and an educator. He pioneered the world’s first electronic recordkeeping software. He served as IBM’s global e-Records Strategy and Business Development Executive. At IBM he was honored as a Technical Leader, one of 439 out of 360,000 IBM employees. Mr. Miller is the recipient of the prestigious Emmett Leahy Award, the highest international recognition given to professionals in the field of information management. His book “Managing Records in Microsoft SharePoint” was an ARMA best seller. Bruce holds a Diploma in Electronics Engineering Technology, an MBA, and is a certified Information Governance Professional.